

Db 481 GTGCTTCCGAGATGGCCACCGCCCTGGCGAAGTGCACGACGAGCTGTTGCTG 540
QY 701 CGGATCTCAAGCTGTGCTGCTTGTCTTGTGCTGACCGGTAGAGAGAAAGCTGTGCTG 760
Db 541 CGGATCTCAAGCTGTGCTGCTTGTCTTGTGCTGACCGGTAGAGAGAAAGCTGTGCTG 600
QY 761 GAGAACCTGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 820
Db 601 GAGAACCTGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 821 GCGTCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 880
Db 661 GCGTCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 881 GAGCGGAGTGTCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 940
Db 721 GAGCGGAGTGTCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
QY 941 TTCAGGAGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1000
Db 781 TTCAGGAGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
QY 1001 CCTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1060
Db 841 CCTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
QY 1061 GCGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1120
Db 901 GCGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
QY 1121 CCCTTACCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1180
Db 961 CCCTTACCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
QY 1181 GCGCTGAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1234
Db 1021 GCGCTGAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1074

RESULT 4
US-09-220-132-135/c
; Sequence 135, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
; APPLICANT: SHYJIAN, Andrew M.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220,132
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 135
; LENGTH: 2559
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-132-135

Query Match 23.9%; Score 505.2; DB 4; Length 2559;
Best Local Similarity 94.6%; Pred. No. 6.8e-121;
Matches 566; Conservative 0; Mismatches 28; Indels 4; Gaps 4;

QY 716 TGTGCTTTGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 775
Db 1634 TGACCTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1575
QY 776 TCTG 835
Db 1574 TCTG 1516

QY 836 GTGGAGCTGAGATCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 895
Db 1515 GTGGAGCTGAGATCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1456
QY 896 AGCTGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 955
Db 1455 AGCTGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1396
QY 956 CCGTCTG 1015
Db 1395 CCGTCTG 1336
QY 1016 GCGCTG 1075
Db 1335 CCGTCTG 1276
QY 1076 GCCACAGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1135
Db 1275 G-CACAGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1218
QY 1136 CGATCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1195
Db 1217 CGATCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1159
QY 1196 AGGAG 1255
Db 1158 AGGAG 1099
QY 1256 TAGCTGCCAAG 1313
Db 1098 TAGCTGCCAAG 1041

RESULT 5
US-08-146-421-4/c
; Sequence 4, Application US/08146421
; Patent No. 513499
; GENERAL INFORMATION:

; APPLICANT: BREMER, GARY
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A POLYPEPTIDE WITH
; TITLE OF INVENTION: ANTI-TUMOR PROPERTIES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESS: 4350 LA JOLLA VILLAGE DRIVE, SUITE 300
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: U.S.A.
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,421
; FILING DATE: 28-OCT-1993
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: PEPPER P.D., FREDERICK W.
; REGISTRATION NUMBER: 31,286
; REFERENCE/DOCKET NUMBER: 489-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-546-4410
; TELEFAX: 619-453-2839
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2562 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA

8/6/96


```

CORRESPONDENCE ADDRESS:
ADDRESS: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HEREMITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1146:
SEQUENCE CHARACTERISTICS:
LENGTH: 1302 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g1480860
US-09-016-434-1146

```

```

Query Match      4.0%; Score 85; DB 4; Length 1302;
Best Local Similarity 48.1%; Pred. No. 2,Se-12;
Matches 241; Conservative 0; Mismatches 260; Indels 0; Gaps 0;

```

```

QY 615 ACCGTATCCCTGAGCTGAGCTGCGCTGCTTCCGCGAGATGCGCACCGCCCTGCGC 674
DB 437 AGCGTTTCCAGTGTGCGACGGCCCGAGGTAATTGTGATGATGACGGCTGGAAT 496
QY 675 ACTGTACCAAGCAGCGGCTGCTGCTGCTGATCTCAAGCTGTGCTTGTCTTGCTG 734
DB 497 ACTGTGATGCGCGGCGATTTGTCAAGGACATCAAGCGGGGAACTGCTCTACCA 556
QY 735 ACCGTAGAGGAAGAAGTGTGCTGAGAACCTTGAGAGACTCTGCGGTGAGCTGGGC 794
DB 557 CCGGTGGCAACCTCAAAATCTCGACCTGGCGCTGCGCGAGGACATGCACTCCGTCG 616
QY 795 CAGATATTCCTCTGTGGAGACAGACGCGTCCGACCTTACGTGGACCTGAGATCTGA 854
DB 617 CGAGCAACACTGCGCGACCAAGCAGCGCTCCCGGCTTCCAGCGCGCCGAGATTCGA 676
QY 855 GCTACGCGGCTCATTAATCTGCGGCAAGCAGCGATGCTGGAGGCTGGCGCTGGCCCT 914
DB 677 ACCGCTTGACACTTCTCTCGGCTTCAAGGTGACATCTGCTGCGTGGAGTACCTCT 736
QY 915 TCAACATGCTGGCGGCACTACCCCTTCAGAGACTCGAGGCTGTCTGCTTTCGGA 974
DB 737 ACAACATCAACAAGGCTGTGACCTCTTGAAGGGGACAAATCTTCAAGTTGTTGAA 796
QY 975 AAGTCCGCGCGGCTTACGCTTTCGCTGACAGGCTCTTGGCGCTTGCCTGCTG 1034
DB 797 ACATCGGGAAGGAGCTACGCGCATCCCGGAGACTGTGGCCCGCTCTGACCTGC 856
QY 1035 TTCGCTGCTCTTGTGGGAGGCAAGCTGAGACGGTCAAGCACAAGCATCTCTTCG 1094
DB 857 TGAAGGATGCTTGAATGAGAACCGGCAAGAGGTTCTTCATTCGGGAGATCCGGCAGC 916

```

```

QY 1095 ACCCTGCTCGACAGGACC 1115
DB 917 ACAGCTGTTCCGGAAGAAC 937

```

```

RESULT 8
US-08-749-902-2
Sequence 2, Application US/08749902
Patent No. 5985635
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goll, Surya K.
APPLICANT: Hillman, Jennifer L.
TITLE OF INVENTION: NOVEL HUMAN SERINE/THREONINE
TITLE OF INVENTION: PROTEIN KINASES
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/749,902
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0150 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-845-4166
TELEFAX: 415-845-0555
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1466 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: Consensus
US-08-749-902-2

```

```

Query Match      3.9%; Score 81.6; DB 2; Length 1466;
Best Local Similarity 47.3%; Pred. No. 2e-11;
Matches 237; Conservative 0; Mismatches 264; Indels 0; Gaps 0;

```

```

QY 615 ACCGTATCCCTGAGCTGAGGCTGCTGCTTCCGCGAGATGCGCACCGCCCTGCGC 674
DB 536 AGCGTTTCCAGTGTGCGACGGCCCGAGGTAATTGTGATGATGACGGCTGGAAT 595
QY 675 ACTGTACCAAGCAGCGTGTGCTGCGGATCTCAAGCTGTGCTTGTCTTGCTGCTG 734
DB 596 ACTGTGATGCGAGGAGGATTTGTGCAAGGCGATCAAGCGGGGAACTGCTGCTACCA 655
QY 735 ACCGTAGAGGAAGAAGCTGTGCTGAGAACCTTGAGAGACTCTTGCCTGACTGAGCG 794
DB 656 CCGGTGGCAACCTCAAAATCTCGACCTGGGCGTGGCGAGGACATGCACTCCGTCG 715
QY 795 CAGATATTCCTCTGTGGAGACAGACGCGTCCGACCTTACGTGGACCTGAGATCTGA 854
DB 716 CGGAGCAACACTGCGGACAGCGCTCCCGGCTTCCAGCGCGCCGAGATTCGA 775
QY 855 GCTACGCGGCTCATTAATCTGCGGCAAGCAGCGATGCTGGAGGCTGGCGCTGGCCCT 914

```

Db 776 ACGGCTGACACTTCTTCGCGCTTCAAGTGAATGATCTGTCGCTGCGGTCACCTCT 835
Qy 915 TCACCATGTGGCCGCGCCACTACCCCTTCAGAGACTCGAGGCTGTCTCTTCGCGCA 974
Db 836 ABAACATCAACAGGGTCTGTACCCCTTGAAGGGGACAACATCTACAGTTGTTGAGA 895
Qy 975 AGATCCGCGCGGGGCTTACGCTTGTGCTGACAGGCTCTCGGCGCCGCTGTCTG 1034
Db 896 ACATCGGGAAGGGAGCTACGCTACCTCCGCGGACTGTGCGCCCGCTCTGACCTGC 955
Qy 1035 TTGCTGCTCTCTTCTGTCGAGGACCAAGCTGAACGCTCACAGCCACAGCATCTCTGC 1094
Db 956 TGAAGGAGATGCTTGAATGACGAACCGGCCAAGAGTTCTTCATCCGCGAGATCCGCGAC 1015
Qy 1095 ACCCTGTGCTGCGACGACCC 1115
Db 1016 ACAAGCTGTTCGGAAGAAC 1036

RESULT 9
US-09-016-434-391

; Sequence 391, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 391:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 288 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAITW03
; CLONE: 2108752
US-09-016-434-391

Query Match 3.4%; Score 71; DB 4; Length 288;
Best Local Similarity 63.3%; Pred. No. 5.6e-09;
Matches 126; Conservative 0; Mismatches 70; Indels 3; Gaps 1;
Qy 806 CTGTGGACACAGACCGGTGCGCCAGCTTACGTGGACCTGAGATCTACAGCTACGCGCC 865

Db 69 CTGAAGACACAGAGGAGGAGCCCTTCTACATCAATCCCGACGTGCTACGCGCGG--- 125
Qy 866 TCATACCTGGGGAAGGACCGCATGTCTGAGGCTGGGCGGCGCTTTCACCATGCTG 925
Db 126 CCGTACCGGTGGACACCGCATGTGAGGCGGCGGCGGTGTCTCTTACCATGCTG 185
Qy 926 GCGGCGCATACCCCTTCAGAGACTCGAGGCTGTCTCTTCTGCGAAGATCCGCGC 985
Db 186 TATGCGCATTCCTCTTACAGACATCCCGAGAGGCTCTTCGGAAGATCAAGGCT 245
Qy 986 GGGGCTTACGCTTGCCTG 1004
Db 246 GCGGATATACCATTCCTG 264

RESULT 10
US-09-016-434-1147

; Sequence 1147, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1147:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2169 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 91488262
US-09-016-434-1147

Query Match 3.1%; Score 65.6; DB 4; Length 2169;
Best Local Similarity 45.9%; Pred. No. 3.2e-07;
Matches 310; Conservative 0; Mismatches 354; Indels 12; Gaps 2;
Qy 448 CAAGGTATACCCCGTCCAGAAACCTGCGGCTGTGAGCCCTTACGCGGCTGCCCC 507
Db 216 CAAGCGCATACGCGGAGAGATCTTAATGATGATGAGCTGACCGAGACCTGAGCA 275
Qy 508 GCACAGCATGTGTGCTGCGCCACATGAGTCTGTGTGTATCCCACTCTCTACGCTT 567

Db 276 CCGGCATGCTGCTGTTTTCGACCACTTGAAGAGCGTGAACAATCAATTTTCTT 335
Qy 568 TTTCACCTGAGCCCATGAGGAGATGACAGCTGTGTGCGAAGCGGCAACCGTATCCCTGA 627
Db 336 GAGAGCTGTGAGCGGAAAGTCCCTGGGCCACATCTGAAGCGCCGGACACCCCTTTTGA 395
Qy 628 GCGTGAAGCTGCGCTGTCTTCCGCGAGATGGCCACCGCCCTGGCGCACTGTCAACGACA 687
Db 396 GCGAGAGTGTGGCTACTACCTGGCGGAGATCCTTTCTGTGCTTCAAGTACTTGGACCGG 455
Qy 688 CGGTGTGTCTGTGTATCTCAAGCTGTGTCTGTGTCTGTGTGTGTGTGTGTGTGTGTGT 747
Db 456 CGGATCTTGCACCGGAGCTCAAGTTGGGAAATTTTTCATCAC-----TGAGACAT 509
Qy 748 GAAGCTGTGTGTGAAGAACCTGAGAGACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 807
Db 510 GGAATGAGAGTGGGAGATTTTGGGCTGGGAGCCCGGTTGAGCCTCGAGGAGAGAGAA 569
Qy 808 GTGGGACAGACAGCGCTGCGGAGCTGAGTGGGAGCTGAGATATCTAGCTCAGCGGCTTC 867
Db 570 GAGAGCATGTGTGGACACCCCACTATGTGTCTCCGAGAGTGTGTGTGTGTGTGTGTGTGT 629
Qy 868 ATACTGGGCGAAGCGAGCCGATGTCTGTGAGCTGTGGGCTGTGGCTGTTCACATGTGCG 927
Db 630 C-----GGCGTGAAGCGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 683
Qy 928 CGGCACTACCCCTTCCAGAGACTCGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 987
Db 684 CGGAGAGCTTCCCTTTGAGAGCGGTGCTGTGAAGAGAACCTACCGCTGTGATCAAGAGGT 743
Qy 988 GGCCTTAGCCTTGTCTGTGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1047
Db 744 TCACTACAGCGCTGT 803
Qy 1048 TCGTGGGAGCGAGCTGTGAAGGCTTCAAGCGCAAGCGATCTCTGTGACCCCTGTGGCG 1107
Db 804 TCGGAGCTCAGCCCGAGAGCGCCCTCTATTGACAGATCTCGGCAATCTTCTTTAC 863
Qy 1108 ACAGAGCCCGATGCCC 1123
Db 864 CAGGCTACACCCCC 879

RESULT 11
US-09-579-664B-4
; Sequence 4, Application US/09579664B
; Patent No. 6514719
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Bird, Timothy A.
; APPLICANT: Varca, G. Duke
; APPLICANT: Martin, Unja
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: NOVEL MORINE AND HUMAN KINASES
; FILE REFERENCE: 2923-A
; CURRENT APPLICATION NUMBER: US/09/579,664B
; CURRENT FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2902
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-579-664B-4

Query Match 2.6%; Score 55.6; DB 4; Length 2902;
Best Local Similarity 49.0%; Pired. No. 0.00014;
Matches 244; Conservative 0; Mismatches 239; Indels 15; Gaps 3;

Qy 578 ACCCATGGGAGATGACAGCGCTGTGGAGCGCGCACCGTATCCCTGAGCTGAGGCT 637
Db 531 AGCCGAGGAGATCTGTATATTATCAATCACTGAGCGGCGCACGCGTGTAGTGGGAGCGCC 590

Qy 638 GCGGTGTCTTTCGCGGAGATGGCCAGCCGCTGTGGCGACTGTACCAAGCAAGTCTGTGTC 697
Db 591 AGGATTTCTTTCGAGAGATGT 650
Qy 698 CTGCGTATCTCAAGCTGT 757
Db 651 CACCGAGATCTCAAGCTGT 710
Qy 758 CTGAGAGCTGTGAGAGACTCTGT 817
Db 711 GACTTGTGCTT-----CTCCAACCTGTACCAAGGCAAGTTCTTCCAGACGTTG 761
Qy 818 CAGCGGTGCGGAGCTTACGAGGAGACTGTGATATCTAGCTCAGGCGCTTACTGTGCGGC 877
Db 762 TGTGGAGCTTCTTCTACGCTGTGCGCTGTGAGATGTGCAAGGAAAGCCC---TATGTGGGC 818
Qy 878 AAGCAGCGAGTGTGTGAGCTGTGGGCGGCGCTTTCACATGTGTGGCGGCGCACTAC 937
Db 819 CAGAGGTGTGACAGCTGT 878
Qy 938 CCTTTCAGGACTGTGAGGCTGT 997
Db 879 CCTTTGACGCGGAGGATCATTAAGACATGTGTGAGCAATCAGTAACGGGCGCTTACCGT 938
Qy 998 TTAGCTGAGGCGCTTGTGGGCGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1057
Db 939 GAGCGCC---CAAGCGTGTCCATGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 995
Qy 1058 CCAGTGAACGCGCTACA 1075
Db 996 CCCAGCGGTGGGCGACA 1013

RESULT 12
US-08-125-468-1/c
; Sequence 1, Application US/08125468
; Patent No. 5589385
; GENERAL INFORMATION:
; APPLICANT: Ryan, Michael J.
; APPLICANT: Lotvin, Jason A.
; APPLICANT: Strathy, Nancy
; APPLICANT: Fancini, Susan E.
; TITLE OF INVENTION: Cloning of the biosynthetic pathway for
; TITLE OF INVENTION: chlorotetracycline and tetracycline Formation and cosmids
; TITLE OF INVENTION: useful therein
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Cyanamid Company
; STREET: One Cyanamid Plaza
; CITY: Wayne
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07470
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/125,468
; FILING DATE: 22-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsevdos, Estelle J
; REGISTRATION NUMBER: 31,145
; REFERENCE/DOCKET NUMBER: 31,255-02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201)831-3241
; TELEFAX: (201)831-3305
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30001 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-125-468-1

Query Match 2.5%; Score 52.8; DB 1; Length 30001;
Best Local Similarity 44.4%; Pred. No. 0.0019;
Matches 213; Conservative 0; Mismatches 267; Indels 0; Gaps 0;

QY 550 CCAGCTCTCTACAGCCTTTTTCACCTGAGACCCATGGGGGACAGACAGCCTGTGGGAC 609
DB 22561 CCAGAGGCTCAACGGGAGTTGGCTTCCGCTCTGGGACGAGCAGCAGAGGAACTCT 22502
QY 610 GGGCCACCGCTATCCCTGAGCCTGAGGCTCCGCTGCTTCCGCCAGATGGCCACCGCCT 669
DB 22501 GCTGGTCCGCGACCGGATGGGCTCAAGCCGCTCTACTACTGCGGACCGCCGCGGCT 22442
QY 670 GGGGCACTGTCAACGACGAGCTGTGCTGCTGCTGATCTCAAGCTGTGTCTTGTCTT 729
DB 22441 GGGCTTCCGCTCGAGCCGACCCAGGCGCTCTGCGGCGACCGGACCTTCCCGCGGCTCG 22382
QY 730 CGCTACCGCTGAGAGGAAAGAGTGTGCTGAGAGAACTGAGAGACTCCTGGTGTGAC 789
DB 22381 CCGGACGGGCTGTGCGAGGTGTGACATGTGAAAGACCCCGAGGCGGCTCTTCTC 22322
QY 790 TGGGCGAGATGATTCCTGTGGGACAGACCGGCTCCAGCCTTACGTGGGAACTTGAGAT 849
DB 22321 CCGGCTCTGCGAGGCTCGCTCCGAGTACGCTCTGCGGCTGGGCGGGGTGGCGG 22262
QY 850 ACTGACCTGACCGGAGCTCTACTGCGGACAGGAGCGGATGTGTGAGAGCTGTGGGCTGGC 909
DB 22261 GAGAGCTGACTGCGGCTGAGAGGCGCGGAGCACACCGACGACTCCGACGACATCGG 22202
QY 910 GCTCTTACCATCTGTGCGGCGGACCTACCTCTTCCAGAGACTCGGAGCTGTGCTCTT 969
DB 22201 CACCGTCCGAGACTGTGTGCGGACACCGTCACTGAGCTGTCTCGAGCTCCGCT 22142
QY 970 CGGCAAGATCCGCGCGGGGCTTACGCGCTTGTGAGGCGCTCTCGGCGCCCTGCGCGCTG 1029
DB 22141 GGGCAGCTGTCTCTCGGCGGCTGTGAGCTGTGCGCGCTGACGCGCTGCGCGCGG 22082

RESULT 13
US-08-474-933-1/C
Sequence 1, Application US/08474933

PATENT No. 5866410
GENERAL INFORMATION:
APPLICANT: Ryan, Michael J.
APPLICANT: Iovlin, Jason A.
APPLICANT: Stralby, Nancy
APPLICANT: Pantini, Susan E.
TITLE OF INVENTION: Cloning of the biosynthetic pathway for
TITLE OF INVENTION: chlorotetracycline and tetracycline formation and cosmids
NUMBER OF INVENTIONS: 1
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSER: American Cyanamid Company
STREET: One Cyanamid Plaza
CITY: Wayne
STATE: New Jersey
COUNTRY: USA
ZIP: 07470

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,933
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/125,468
FILING DATE: 22-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Tsevdos, Estelle J
REGISTRATION NUMBER: 31,145
REFERENCE/DOCKET NUMBER: 31,255-02
TELEPHONE: (201)831-3241
TELEFAX: (201)831-3305
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 30001 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-474-933-1

Query Match 2.5%; Score 52.8; DB 2; Length 30001;
Best Local Similarity 44.4%; Pred. No. 0.0019;
Matches 213; Conservative 0; Mismatches 267; Indels 0; Gaps 0;

QY 550 CCAGCTCTCTACAGCCTTTTTCACCTGAGACCCATGGGGGACATGACAGCCTGTGGGAC 609
DB 22561 CCAGAGGCTCAACGGGAGTTGGCTTCCGCTCTGGGACGAGCAGCAGAGGAACTCT 22502
QY 610 GGGCCACCGCTATCCCTGAGCCTGAGGCTCCGCTGCTTCCGCCAGATGGCCACCGCCT 669
DB 22501 GCTGGTCCGCGACCGGATGGGCTCAAGCCGCTCTACTACTGCGGACCGCCGCGGCT 22442
QY 670 GGGGCACTGTCAACGACGAGCTGTGCTGCTGCTGATCTCAAGCTGTGTCTTGTCTT 729
DB 22441 GGGCTTCCGCTCGAGCCGACCCAGGCGCTCTGCGGCGACCGGACCTTCCCGCGGCTCG 22382
QY 730 CGCTACCGCTGAGAGGAAAGAGTGTGCTGAGAGAACTGAGAGACTCTGCGTGTGAC 789
DB 22381 CCGGACGGGCTGTGCGAGGTGTGACATGTGAAAGACCCCGAGGCGGCTCTTCTC 22322
QY 790 TGGGCGAGATGATTCCTGTGGGACAGACCGGCTCCAGCCTTACGTGGGAACTTGAGAT 849
DB 22321 CCGGCTCTGCGAGGCTCGCTCCGAGTACGCTCTGCGGCTGGGCGGGGTGGCGG 22262
QY 850 ACTGACCTGACCGGAGCTCTACTGCGGACAGGAGCGGATGTGTGAGAGCTGTGGGCTGGC 909
DB 22261 GAGAGCTGACTGCGGCTGAGAGGCGCGGAGCACACCGAGACTGTGCCAGACATCGG 22202
QY 910 GCTCTTACCATCTGTGCGGCGGACCTACCTCTTCCAGAGACTCGGAGCTGTGCTCTT 969
DB 22201 CACCGTCCGAGACTGTGTGCGGACACCGTCACTGAGCTGTCTCGAGCTCCGCT 22142
QY 970 CGGCAAGATCCGCGCGGGGCTTACGCGCTTGTGAGGCGCTCTCGGCGCCCTGCGCGCTG 1029
DB 22141 GGGCAGCTGTCTCTCGGCGGCTGTGAGCTGTGCGCGCTGACGCGCTGCGCGCGG 22082

RESULT 14
US-09-691-861A-14
Sequence 14, Application US/09691861A

PATENT No. 6482935
GENERAL INFORMATION:
APPLICANT: Wei, Ming-Hui et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: CLO00892
CURRENT APPLICATION NUMBER: US/09/691,861A
CURRENT FILING DATE: 2000-10-18
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 14
LENGTH: 601
TYPE: DNA
ORGANISM: Homo sapiens

FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(601)
; OTHER INFORMATION: n = A,T,C or G
US-09-691-861A-14

Query Match
Best Local Similarity 49.7%; Score 51; DB 4; Length 601;
Matches 188; Conservative 0; Mismatches 181; Indels 9; Gaps 2;

QY 580 CCATGGGACATGACAGCCTGAGGAAACCGGACCGGATCCCTGAGCCTGAGGCTGC 639
DB 73 CCAAGGCGACCTCTCGAGTTAATCAAAACCGGGAGCCCTGATGAGACGAAAGCTCG 132
QY 640 CGTGCTCTTCGCGCATGAGCCGCGCTGCGCATGTCACGACGAGCTGAGTCT 699
DB 133 CAAAGAGTTCCACGAGCTTCTTGCGCATGAGTACGACGACCTGAGAGCTGCTCA 192
QY 700 GCGTATCTCAAGCTGAGTCTGCTTTGCTTCTGAGACCGTGAAGAAAGAGTGTCT 759
DB 193 CCGGAGACTCAAGGTGACCAACCTTCTGACAGAGACTTCAACATCAAGCTGTCCGA 252
QY 760 GAGAACCTGAGAGACTCTGCTGCTGATGAGGCGACAGATATCCCTGAGGACAAAGCA 819
DB 253 CTTCAGCTTCTCCCAAGGCTGCTGCGGATGACAGTGTGAAATGCGNTTAAGCAAGAC 312
QY 820 C-----GCGTCCGAGCTTACGTGAGACCTGAGACTCAGCTACAGGCGCTCATCTC 873
DB 313 CTTCGTGGGTGACACGATGAGGCGGCGGAGAGTGT---GCAAGGCAATCCCTACCA 369
QY 874 GGGGAGGACCGCATGATGTTGAGACCTGAGGCTGAGGCTCTTCAACATGCTGCGCGCA 933
DB 370 GCCCAAGGTGACGATCTGAGACCTGAGGCTGAGTATCTTACATCATGCTGCGGCTC 429
QY 934 CTACCCCTTCAGGACTC 951
DB 430 CATGCCCTACGAGACTC 447

RESULT 15
US-09-691-861A-1
; Sequence 1, Application US/09691861A
; Patent No. 6482935
; GENERAL INFORMATION:
; APPLICANT: Wei, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: C1000892
; CURRENT APPLICATION NUMBER: US/09/691, 861A
; CURRENT FILING DATE: 2000-10-18
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1104
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-691-861A-1

Query Match
Best Local Similarity 49.7%; Score 50.4; DB 4; Length 1104;
Matches 188; Conservative 0; Mismatches 181; Indels 9; Gaps 2;

QY 580 CCATGGGACATGACAGCCTGAGGAAACCGGACCGGATCCCTGAGCCTGAGGCTGC 639
DB 282 CCAAGGCGACCTCTCGAGTTAATCAAAACCGGGAGCCCTGATGAGACGAAAGCTCG 341
QY 640 CGTGCTCTTCGCGCATGAGCCGCGCTGCGCATGTCACGACGAGTCTGAGTCT 699
DB 342 CAAAGAGTTCCACGAGCTTCTTGCGCATGAGTACGACGACCTGAGAGCTGCTCA 401
QY 700 GCGTATCTCAAGCTGAGTCTGCTTTGCTTCTGAGACCGTGAAGAAAGCTGTCT 759

DB 402 CCGGAGCTCAAGGTGACCAACCTTCTTGAACAAGACTTCAACATCAAGCTGTCCGA 461
QY 760 GAGAACCTGAGAGACTCTCTGCTGCTGATGAGGCGAGATATTCCTGTGGACAAAGCA 819
DB 462 CTTCAGCTTCTCCAAAGGCTGCTGCGGATGACAGTGTGAATGATTAAGCAAGAC 521
QY 820 C-----GCGTCCGAGCTTACGTGAGACCTGAGATCTCAAGCTGAGGCGCTCATACTC 873
DB 522 CTTCGTGGGTGACACGATGAGGCGGCGGCGGAGGTGT---GCAAGGCAATTCCTACCA 578
QY 874 GGGGAGGACCGCATGATGTTGAGACCTGAGGCTGAGGCTCTTCAACATGCTGAGCGGCA 933
DB 579 GCCCAAGGTGACGATCTGAGACCTGAGGCTGATCTCTACATCATGCTGTGCGGCTC 638
QY 934 CTACCCCTTCAGGACTC 951
DB 639 CATGCCCTACGAGACTC 656

Search completed: August 29, 2004, 12:06:29
Job time : 159 secs

QY 1 GGAGGCGGCTCCGCGCGCTCCGCTGCTAGAGACCCGGGACAGGGCTGAGAGCTGGGCTGGGA 60
Ddb 222 GGAAGCGGCTCCGCGCGCTCCGCTGCTAGAGACCCGGGACAGGGCTGAGAGCTGGGCTGGGA 280
QY 61 TCCGAGAGCTCCGAGAGCGACA-CGGACCGGCCACCTGCTGGTGCCTTGGAGGCTCTGA 115

Db 282 TCCCAAGCTCCGACAGAGCGAGCGGCGGCGCCCACTGCTGTGCTCCCTGAGAGGCTCTGA 341
 QY 120 GCCCGGCGGCGCGCGGCGCCCAAGGAGAACGAGCGGAGATGCGAGCACTCCCTTGG 179
 Db 342 GCCCGGCGGCGCGCGGCGCCCAAGGAGAACGAGCGGAGATGCGAGCACTCCCTTGG 401
 QY 180 CTGCTCTGCGGGTTCCTGTCTCAGAGAAAGAGGTTGAGATGATGACAACTTAGTA 239
 Db 402 CTGCTCTGCGGGTTCCTGTCTCAGAGAAAGAGGTTGAGATGATGACAACTTAGTA 461
 QY 240 CGAGCGCTCCGCTCCAGAAAGAGTCCGAGTGGGCCCCAGCCGAGCTGCCCCCTGACC 299
 Db 462 CGAGCGCTCCGCTCCAGAAAGAGTCCGAGTGGGCCCCAGCCGAGCTGCCCCCTGACC 521
 QY 300 TGTTCGCCCTGAGCCCACTACTGTCCAGATCCGAGCACTGCTGTGGCCCTGCTCCC 359
 Db 522 TGTTCGCCCTGAGCCCACTACTGTCCAGATCCGAGCACTGCTGTGGCCCTGCTCCC 581
 QY 360 GTCTTGGGCGCTATGTCTCTGAGAGCCGAGAGGCGGCGGCGCTTACAGAGCCCTGAC 419
 Db 582 GTCTTGGGCGCTATGTCTCTGAGAGCCGAGAGGCGGCGGCGCTTACAGAGCCCTGAC 641
 QY 420 ACTGCGCTTACAGAGCACTGATATCTTGCAGAGGTTACCCCGTCCAGAGACCTTGGCCG 479
 Db 642 ACTGCGCTTACAGAGCACTGATATCTTGCAGAGGTTACCCCGTCCAGAGACCTTGGCCG 701
 QY 480 TGTGAGAGCCCTTACAGAGCGGCGCTGCCCCGCAAGAGCACTGAGCTGGGCCCCACTGAGTCC 539
 Db 702 TGTGAGAGCCCTTACAGAGCGGCGCTGCCCCGCAAGAGCACTGAGTGGGCCCCACTGAGTCC 761
 QY 540 TGTGCTGATCCCAAGCTCTCTTACGCGCTTTTTCATCTGAGACCCATGAGGAGATGACAGCC 599
 Db 762 TGTGCTGATCCCAAGCTCTCTTACGCGCTTTTTCATCTGAGACCCATGAGGAGATGACAGCC 821
 QY 600 TGTGAGCAAGCGCGCACTGATCCCTGAGGCTGAGGCTGCGCTGCTTTCCGCGCAGATGG 659
 Db 822 TGTGAGCAAGCGCGCACTGATCCCTGAGGCTGAGGCTGCGCTGCTTTCCGCGCAGATGG 881
 QY 660 CCAAGCGCGCTGAGCGCACTGTCAACAGACAGGCTGAGTCTGCGGATCTCAAGCTGTGTC 719
 Db 882 CCAAGCGCGCTGAGCGCACTGTCAACAGACAGGCTGAGTCTGCGGATCTCAAGCTGTGTC 941
 QY 720 GCTTTGTCTTCTGCTGACCTGAGAGAGAAAGCTGTGCTGAGAACTTGAAGACTCTCT 779
 Db 942 GCTTTGTCTTCTGCTGACCTGAGAGAGAAAGCTGTGCTGAGAACTTGAAGACTCTCT 1001
 QY 780 GCGTGTGATCTGGGCGCAGATGATTCCTGTGAGCAAGCACTGCTCCAGCTTACGTGG 839
 Db 1002 GCGTGTGATCTGGGCGCAGATGATTCCTGTGAGCAAGCACTGCTCCAGCTTACGTGG 1061
 QY 840 GACTGTGATCTGACACTGACGAGGCTCTCACTGAGGCAAGGAGCGAGTGTGAGAGCC 899
 Db 1062 GACTGTGATCTGACACTGACGAGGCTCTCACTGAGGCAAGGAGCGAGTGTGAGAGCC 1121
 QY 900 TGGGCGTGGGCGCTCTTCAACATGTGTGGCGGCACTTACCTTTTCAAGAACTGAGAGCTTG 959
 Db 1122 TGGGCGTGGGCGCTCTTCAACATGTGTGGCGGCACTTACCTTTTCAAGAACTGAGAGCTTG 1181
 QY 960 TCTGTCTTCTGCGCAAGATCCGCGGCGGAGGCTTGAAGCTTGTGAGAGGCTCTGAGGCC 1019
 Db 1182 TCTGTCTTCTGCGCAAGATCCGCGGCGGAGGCTTGAAGCTTGTGAGAGGCTCTGAGGCC 1241
 QY 1020 CTGCGCGCTGTCTGATTCGCTGCTCTCTTCTGTGAGAGCCAGCTGAAACGGCTCAAGCCA 1079
 Db 1242 CTGCGCGCTGTCTGATTCGCTGCTCTCTTCTGTGAGAGCCAGCTGAAACGGCTCAAGCCA 1301
 QY 1080 CAGGAGATCTCTCTGACACCTTGTGCTGCGACAGAGACCCGATGCTTGAAGCCCAACCCGAT 1139
 Db 1302 CAGGAGATCTCTCTGACACCTTGTGCTGCGACAGAGACCCGATGCTTGAAGCCCAACCCGAT 1361
 QY 1140 CCAATCTCTGAGAGGCTGCGCAGGTGCTCCCTGATGAGCTGAGGCTGAGCAAGACCAAGG 1199

Db 1362 CCAATCTCTGAGAGGCTGCGCAGGTGCTCCCTGATGAGCTGAGGCTGAGCAAGACCAAGG 1421
 QY 1200 AAGAGAGGAGACAGAGATGTTCTGTATGCTTGAAGACCAACCTTACTAGAGCTGAG 1259
 Db 1422 AAGAGAGGAGACAGAGATGTTCTGTATGCTTGAAGACCAACCTTACTAGAGCTGAG 1481
 QY 1260 CTGCGCAAGATGATGATGATTTGGGCGGTAGCTCCAGGCTTCTCTGCTGCTGAACTGAG 1319
 Db 1482 CTGCGCAAGATGATGATGATTTGGGCGGTAGCTCCAGGCTTCTCTGCTGCTGAACTGAG 1541
 QY 1320 CCAACCTTCTAGTCTCTCCAGAGGAGAGAGAGAGAGGCTGATGAGAGTGTGCTGTG 1379
 Db 1542 CCAACCTTCTAGTCTCTCCAGAGAGAGAGAGAGAGGCTGATGAGAGTGTGCTGTG 1601
 QY 1380 TACACATCTGTTTGTTCACACACATGCAAGTTCCTGCTGGGTCTTATGAGGTCGAA 1439
 Db 1602 TACACATCTGTTTGTTCACACACATGCAAGTTCCTGCTGGGTCTTATGAGGTCGAA 1661
 QY 1440 GCCCTGTTCTGGGCTGAGAGTACAGCATGAGAGCAATATTCCTGCTGAC 1499
 Db 1662 GCCCTGTTCTGGGCTGAGAGTACAGCATGAGAGCAATATTCCTGCTGAC 1721
 QY 1500 AGAGATGACAAACTGGGATCCTTGAAGCTGACCAACCTTTTCAAGACATGAGTCACTGT 1559
 Db 1722 AGAGATGACAAACTGGGATCCTTGAAGCTGACCAACCTTTTCAAGACATGAGTCACTGT 1781
 QY 1560 CTACATCTGGGTACCTTGTACAGATGTGGGCTTCCATGATGCTGTGCTCAGGCACT 1619
 Db 1782 CTACATCTGGGTACCTTGTACAGATGTGGGCTTCCATGATGCTGTGCTCAGGCACT 1841
 QY 1620 CTGTCCAGAGCAATCCCTTTTCAAAACCAACAGAGCTCTTGTATCTTGTATCTTTTC 1679
 Db 1842 CTGTCCAGAGCAATCCCTTTTCAAAACCAACAGAGCTCTTGTATCTTGTATCTTTTC 1901
 QY 1680 AGAGAAAGGAGGTATCCCTGTGCTCAAGGCTCCAGGCTCTCCCGCGCAACTCAGGACC 1739
 Db 1902 AGAGAAAGGAGGTATCCCTGTGCTCAAGGCTCCAGGCTCTCCCGCGCAACTCAGGACC 1961
 QY 1740 CAAAGCCAGCTCACTCTGGGAACTGTGTTCCAGCATCTCTGTCTCTTGAATTAAAGAT 1799
 Db 1962 CAAAGCCAGCTCACTCTGGGAACTGTGTTCCAGCATCTCTGTCTCTTGAATTAAAGAT 2021
 QY 1800 TCTCTTCCAGGCTTGAAGCTTGGGAGTTCAGGAGTAAAGATTCAACTTGAAGCT 1859
 Db 2022 TCTCTTCCAGGCTTGAAGCTTGGGAGTTCAGGAGTAAAGATTCAACTTGAAGCT 2081
 QY 1860 AGTTCTGTCTTACTCAAGTGTCTTGAATGAGGCTCAGGCTGTCAACCAATGGGCT 1919
 Db 2082 AGTTCTGTCTTACTCAAGTGTCTTGAATGAGGCTCAGGCTGTCAACCAATGGGCT 2141
 QY 1920 TTCTGACTGAGACCAAGTGTGAGGACAGATTTAGGAGGCTGTCTGTGAGGCAAC 1979
 Db 2142 TTCTGACTGAGACCAAGTGTGAGGACAGATTTAGGAGGCTGTCTGTGAGGCAAC 2201
 QY 1980 TGAAGAGTCCAGGTGGAGCTCTTCTGAGGAGCACTTGGGCTTCAACTCCAGGCTCAAT 2039
 Db 2202 TGAAGAGTCCAGGTGGAGCTCTTCTGAGGAGCACTTGGGCTTCAACTCCAGGCTCAAT 2261
 QY 2040 CTCTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2099
 Db 2262 CTCTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2321
 QY 2100 GAAATTAATAAAAAA 2116
 Db 2322 GAAATTAATAAAAAA 2338

RESULT 2
 US-10-649-156-7
 ; Sequence 7, Application US/10649156
 ; Publication No. US20040038346A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Meyers, Rachel

```
/ APPLICANT: Kapeller-Libermann, Rosana
/ APPLICANT: Williamson, Mark
/ TITLE OF INVENTION: No. US20040038346A1el Human Protein Kinases and Uses
/ TITLE OF INVENTION: Therefor
/ FILE REFERENCE: 35800/209996
/ CURRENT APPLICATION NUMBER: US/10/649,156
/ CURRENT FILING DATE: 2003-08-27
/ PRIOR APPLICATION NUMBER: US/09/799,875
/ PRIOR FILING DATE: 2001-03-06
/ PRIOR APPLICATION NUMBER: 60/182,059
/ PRIOR FILING DATE: 2000-02-11
/ PRIOR APPLICATION NUMBER: 09/659,287
/ PRIOR FILING DATE: 2000-09-12
/ NUMBER OF SEQ ID NOS: 32
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 7
/ LENGTH: 2389
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (383)...(1456)
/ US-10-649-156-7

Query Match      99.3%; Score 2100.2; DB 13; Length 2389;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2113; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
```

```

Db      882  CCACCGCCCTGGCCACCTGACACGACCGGTGCTCCGTGATCTCAAGCTGTGTC 941
Qy      720  GCTTGTCTTCTGCTGACCGGTGAGGAAAGAGCTGTGCTGAGAACTTGAGAGACTCT 779
Db      942  GCTTGTCTTCTGCTGACCGGTGAGGAAAGAGCTGTGCTGAGAACTTGAGAGACTCT 1001
Qy      780  GCGTGTCTGCTGAGGCGGAGATATCCCTGTGGGCAAGACGCGGTCCAGCTTACGTG 839
Db      1002  GCGTGTCTGCTGAGGCGGAGATATCCCTGTGGGCAAGACGCGGTCCAGCTTACGTG 1061
Qy      840  GACCTGAGATCTGACCTCAGCGGCTTATCTGCGGCAAGGCAAGCTGATGTCTGAGCC 899
Db      1062  GACCTGAGATCTGACCTCAGCGGCTTATCTGCGGCAAGGCAAGCTGATGTCTGAGCC 1121
Qy      900  TGGGCGTGGCGCTTTCACCATGCTGCGCGGCACTACCTCTTCCAGGACTCGGAGCC 959
Db      1122  TGGGCGTGGCGCTTTCACCATGCTGCGCGGCACTACCTCTTCCAGGACTCGGAGCC 1181
Qy      960  TCTGTCTCTTGGGCAAGATCCGCGGCGGAGGCTTACGCTTGGCTCAGGCTCTGAGCC 1019
Db      1182  TCTGTCTCTTGGGCAAGATCCGCGGCGGAGGCTTACGCTTGGCTCAGGCTCTGAGCC 1241
Qy      1020  CTGCGCGCTGTGCTGCTGCTGCTCTTCTGCTGCGGAGCCAGCTGAACTGCTACAGCA 1079
Db      1242  CTGCGCGCTGTGCTGCTGCTGCTCTTCTGCTGCGGAGCCAGCTGAACTGCTACAGCA 1301
Qy      1080  CAGGCACTCTCTGCAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1139
Db      1302  CAGGCACTCTCTGCAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1361
Qy      1140  CCGATCTCTGAGGAGCTGCGGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1199
Db      1362  CCGATCTCTGAGGAGCTGCGGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1421
Qy      1200  AAGAGAGGAGAGACAGAGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1259
Db      1422  AAGAGAGGAGAGACAGAGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1481
Qy      1260  CTGCGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1319
Db      1482  CTGCGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1541
Qy      1320  CCAGACCTTCAAGGCTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1379
Db      1542  CCAGACCTTCAAGGCTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1601
Qy      1380  TACACATCTGCTTGTTCACACATGACATGACATGACATGACATGACATGACATGAC 1439
Db      1602  TACACATCTGCTTGTTCACACATGACATGACATGACATGACATGACATGACATGAC 1661
Qy      1440  GCCCTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1499
Db      1662  GCCCTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1721
Qy      1500  AGAGATGACAACTGGATCTTGAAGTGAACAACTTTCATGACATGACATGACATGAC 1559
Db      1722  AGAGATGACAACTGGATCTTGAAGTGAACAACTTTCATGACATGACATGACATGAC 1781
Qy      1560  CTACACTGAGTACACTTGTACAGATGCTGAGCTTCACTGATGCTGCTGCTGCTGCT 1619
Db      1782  CTACACTGAGTACACTTGTACAGATGCTGAGCTTCACTGATGCTGCTGCTGCTGCT 1841
Qy      1620  CTGTCCAGAGCAATCCCTTTCACAAACAAACAAACAAACAAACAAACAAACAAAC 1679
Db      1842  CTGTCCAGAGCAATCCCTTTCACAAACAAACAAACAAACAAACAAACAAACAAAC 1901
Qy      1680  AGAGAAAGGAGGATATCCCTGCTGAGGCTTCCAGGCTTCCCTGCAACTCAGAGACC 1739
Db      1902  AGAGAAAGGAGGATATCCCTGCTGAGGCTTCCAGGCTTCCCTGCAACTCAGAGACC 1961
Qy      1740  CAAAGCCAGCTCACTGTGGAACTGTGTCCAGCATCTGTCTCTTGAATGAAGAT 1799
```

Db 1962 CAAGCCAGCTCCTCTGAGGAAGTGTGTTCCAGCATCTCTGCTCTGATTAGAGAT 2021
 Qy 1800 TCTCTTCAGGCTTAAGCTGTGAGATTGGGCGCAGATTAAGATCCAACTATAGGCT 1859
 Db 2022 TCTCTTCAGGCTTAAGCTGTGAGATTGGGCGCAGATTAAGATCCAACTATAGGCT 2081
 Qy 1860 AGTTCTTGTCTTACTCAAGACTGTTCTGAAATGAGGGTCCAGGCTGTCAACATGAGGC 1919
 Db 2082 AGTTCTTGTCTTACTCAAGACTGTTCTGAAATGAGGGTCCAGGCTGTCAACATGAGGC 2141
 Qy 1920 TTCGACCTGAGCAGCAGATGAGGGGACAGGATTAAGCAGGCTGTCTGTGCGCAC 1979
 Db 2142 TTCGACCTGAGCAGCAGATGAGGGGACAGGATTAAGCAGGCTGTCTGTGCGCAC 2201
 Qy 1980 TGAAGAGTCCAGAGTGGAGCTTCTGAGGACACTTGGGGTCCCAATCCAGATCCATA 2039
 Db 2202 TGAAGAGTCCAGAGTGGAGCTTCTGAGGACACTTGGGGTCCCAATCCAGATCCATA 2261
 Qy 2040 CTCTAGGTTTGGATACCATGATGATGATGATGATGATGATGATGATGATGATGAT 2099
 Db 2262 CTCTAGGTTTGGATACCATGATGATGATGATGATGATGATGATGATGATGATGAT 2321
 Qy 2100 GAAATTAATAAAAAAAAAA 2116
 Db 2322 GAAATTAATAAAAAAAAAA 2338

RESULT 3
 US-10-425-114-18433
 ; Sequence 18433, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jindong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E.
 ; APPLICANT: Tabaska, Jack E.
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(5311)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; NUMBER OF SEQ ID NOS: 2003-04-28
 ; SEQ ID NO 18433
 ; LENGTH: 2082
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: LIB3082-008-C10_FLI
 ; US-10-425-114-18433

Query Match 97.4%; Score 2062; DB 13; Length 2082;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 2076; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 24 CTGCTAGAGACCGGAGGAGGCTGAGAGCTGGGATCCCGAGCTCCGAGAGAGGAG- 82
 Db 1 CTGCTAGAGACCGGAGGAGGCTGAGAGCTGGGATCCCGAGCTCCGAGAGAGGAG 60
 Qy 83 CGGGCGGCGCCACCTGCTGTGTGCTCTGAGGCTCTGAGGCTCCGAGCGCGCGCGGCCAC 142
 Db 61 CGGGCGGCGCCACCTGCTGTGTGCTCTGAGGCTCTGAGGCTCCGAGCGCGCGCGGCCAC 120
 Qy 143 GCGGAACGACGCGGAGAGTGGAGACACCCCTGTGGCTGTCTGTCTGTCTGTCTGTCTC 202
 Db 121 GCGGAACGACGCGGAGAGTGGAGACACCCCTGTGGCTGTCTGTCTGTCTGTCTGTCTC 180
 Qy 203 AGGAAGAAAGCGGTTGAGTTGATGACAACTTAAGATACGAGAGCTCCCGTCAGAAAGA 262
 Db 181 AGGAAGAAAGCGGTTGAGTTGATGACAACTTAAGATACGAGAGCTCCCGTCAGAAAGA 240
 Qy 263 GCTGAGAGTGGGCGCCAGCCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 322

Db 241 GCTGGAAGTGGGCGCCAGCCAGACTGCCCCCTGCTGTGTGCTGCTGCTGCTGCTGCTGCT 300
 Qy 323 GCTCAGATCGTNGCAACTGCTGTGGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 382
 Db 301 GCTCAGATCGTNGCAACTGCTGTGGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 Qy 383 GAGCCCGAGGAGGCGGCGGCGCTTACAGGCGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCT 442
 Db 361 GAGCCCGAGGAGGCGGCGGCGCTTACAGGCGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCT 420
 Qy 443 ACCGGAAGGTGATACCCCTTCAAGAAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 502
 Db 421 ACCGGAAGGTGATACCCCTTCAAGAAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 Qy 503 CCCCCGACAGCATGTGTGCTGCGCCCACTGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 562
 Db 481 CCCCCGACAGCATGTGTGCTGCGCCCACTGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 Qy 563 GCTTTTTCATCTGAGACCATGAGGACATGACAGCTGTGTGAGAGCGGCGCACCTGATC 622
 Db 541 GCTTTTTCATCTGAGACCATGAGGACATGACAGCTGTGTGAGAGCGGCGCACCTGATC 600
 Qy 623 CTTGAGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 682
 Db 601 CTTGAGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 Qy 683 CAGCAGGCTGAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 742
 Db 661 CAGCAGGCTGAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 Qy 743 AGGAAGAGTGTGTGCTGAGAACTGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 802
 Db 721 AGGAAGAGTGTGTGCTGAGAACTGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 Qy 803 TCCCTGTGGGACAAAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 862
 Db 781 TCCCTGTGGGACAAAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 Qy 863 GCCTCTACTCTGGGCAAGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 922
 Db 841 GCCTCTACTCTGGGCAAGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 Qy 923 CTGGCGGCGCACTACCCCTTCCAGAGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 982
 Db 901 CTGGCGGCGCACTACCCCTTCCAGAGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 Qy 983 CGGCGGCGCTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1042
 Db 961 CGGCGGCGCTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 Qy 1043 CTCTTGTGTGAGAGCAGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAG 1102
 Db 1021 CTCTTGTGTGAGAGCAGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAG 1080
 Qy 1103 CTGCGACAGAACCGATGCTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1162
 Db 1081 CTGCGACAGAACCGATGCTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
 Qy 1163 GTGTGCT 1222
 Db 1141 GTGTGCT 1200
 Qy 1223 GTTCTGTATGCTTGAAGACACCTTACTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1282
 Db 1201 GTTCTGTATGCTTGAAGACACCTTACTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
 Qy 1283 GGGGTAGCTCCAGGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1342
 Db 1261 GGGGTAGCTCCAGGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
 Qy 1343 AGGGAAGAGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1402

Db	961	CTGTGAGGGCTCTCGGGCCCTGCGCCGCTGTCTGTGTCGTCGCTCTCTTGTGTGGAGCCA	1020
QY	1061	GTGTGAACGGCTCAAGACCAAGGCAATCTCTTGCACACCTCTGGCTCGACAGGACCCGATG	1120
Db	1021	GTGTGAACGGCTCAAGACCAAGGCAATCTCTTGCACACCTCTGGCTCGACAGGACCCGATG	1080
QY	1121	CCCTTGGCCCAACCCGATCCCATCTCTGGAGGGCGCCCAAGTGTCCCTGATGGAATG	1180
Db	1081	CCCTTGGCTCGAACCCGATCCCATCTCTGGAGGGCGCCCAAGTGTCCCTGATGGAATG	1140
QY	1181	GGGCTTGGAGGAAGCCAGGGAAGAGAGGGAGACAGAGAAATGTGTCTGTATGGCTTAGAC	1240
Db	1141	GGGCTTGGAGGAAGCCAGGGAAGAGAGGGAGACAGAGAAATGTGTCTGTATGGCTTAGAC	1200
QY	1241	CACCTTACTACACGCTCAGCTGACCAACAGTGGATTAAGTTTGGGGGTAGCTTCCAGCTT	1300
Db	1201	CACCTTACTACACGCTCAGCTGACCAACAGTGGATTAAGTTTGGGGGTAGCTTCCAGCTT	1260
QY	1301	CTCTTCCCTCTGAACCTGAGCCAAACCTTCAAGCCCTTCAGAGGAGGAAGAAAGCAGAAAGC	1360
Db	1261	CTCTTCCCTCTGAACCTGAGCCAAACCTTCAAGCCCTTCAGAGGAGGAAGAAAGCAGAAAGC	1320
QY	1361	CTGTGGAGTGGTGGTGGTGAACATCTGTCTTGTTCACACACATGCAATTCCTGCTTG	1420
Db	1321	CTGTGGAGTGGTGGTGGTGAACATCTGTCTTGTTCACACACATGCAATTCCTGCTTG	1380
QY	1421	GGTGTCTTACAGTGCCAAGCCCTGTCTCGGTGTGGGAGTACAGCAGTAGCAAAAGGA	1480
Db	1381	GGTGTCTTACAGTGCCAAGCCCTGTCTCGGTGTGGGAGTACAGCAGTAGCAAAAGGA	1440
QY	1481	GACCAATATTCCTGTCTCACAGAGATGACAACTGGCAATCCTTGAAGCTGACAAACTTTTC	1540
Db	1441	GACCAATATTCCTGTCTCACAGAGATGACAACTGGCAATCCTTGAAGCTGACAAACTTTTC	1500
QY	1541	CATGACCAATAGGTCACTGTCTTACACTGGGGTACACTTGTGACAGTGTGGCTTCCATGA	1600
Db	1501	CATGACCAATAGGTCACTGTCTTACACTGGGGTACACTTGTGACAGTGTGGCTTCCATGA	1560
QY	1601	TGCTGTGTCTCAAGCACTGTGTCCAAAGACAATCCCTTTCACAAACAAACCAAGCTTGCCT	1660
Db	1561	TGCTGTGTGTCTCAAGCACTGTGTCCAAAGACAATCCCTTTCACAAACAAACCAAGCTTGCCT	1620
QY	1661	TTGTATCTTGTACCTTTTACAGAAAGAGAGAGTATCCCTGTGCAAAAGGCTTCCAGGCTTC	1720
Db	1621	TTGTATCTTGTATCCTTTTACAGAAAGAGAGAGTATCCCTGTGCAAAAGGCTTCCAGGCTTC	1680
QY	1721	TCCCTTGCACACTCAGAACCCAAAGCCCAAGCTCACTGTGGAACTGTGTGCCAGACTCTCT	1780
Db	1681	TCCCTTGCACACTCAGAACCCAAAGCCCAAGCTCACTGTGGAACTGTGTGCCAGACTCTCT	1740
QY	1781	GTCTCTTGTATTAAGAGATTTCTCTTCCAGGCTTAAGCTTGGGATTTTGGGCGAGATTA	1840
Db	1741	GTCTCTTGTATTAAGAGATTTCTCTTCCAGGCTTAAGCTTGGGATTTTGGGCGAGATTA	1800
QY	1841	GAATCCAAACTATAGGCTATGTTCTTGTCTTCACTCAAGACTGTGTGGAATGAGGGTCCA	1900
Db	1801	GAATCCAAACTATAGGCTATGTTCTTGTCTTCACTCAAGACTGTGTGGAATGAGGGTCCA	1860
QY	1901	GGCTGTTCACATGAGGGCTTTTGAACCTGAGCAACAAAGTTGAGGAGACAGATTAAGGAG	1960
Db	1861	GGCTGTTCACATGAGGGCTTTTGAACCTGAGCAACAAAGTTGAGGAGACAGATTAAGGAG	1920
QY	1961	GGTGTGTCTGTGGCCCACTGGAAAGTCCAGAGTGGAGACTTTTGGGGACACTTGGGGGT	2020
Db	1921	GGTGTGTCTGTGGCCCACTGGAAAGTCCAGAGTGGAGACTTTTGGGGACACTTGGGGGT	1980
QY	2021	CCACAAATCCCAAGGTCCATTACTTAAGTTTGGATATACATGATGATATGTTTACCTGTG	2080
Db	1981	CCACAAATCCCAAGGTCCATTACTTAAGTTTGGATATACATGATGATATGTTTACCTGTG	2040
QY	2081	CCCAATTAAGAGAAATTAATGAATTA 2106	
Db	2041	CCCAATTAAGAGAAATTAATGAATTA 2066	

	RESULT 5	US-10-098-841-271
/	Sequence 271, Application US/10098841	
/	Publication No. US20020197679A1	
/	GENERAL INFORMATION:	
/	APPLICANT: Tang, Y. Tom	
/	APPLICANT: Liu, Chenghua	
/	APPLICANT: Asundi, Vinod	
/	APPLICANT: Xu, Chongjun	
/	APPLICANT: Zhou, Ping	
/	APPLICANT: Ma, Yunting	
/	APPLICANT: Wang, Jian-Rui	
/	APPLICANT: Zhao, Qing A.	
/	APPLICANT: Ren, Feiyang	
/	APPLICANT: Chen, Rui-hong	
/	APPLICANT: Wang, Dunrui	
/	APPLICANT: Wang, Zhiwei	
/	APPLICANT: Mehmman, Tom	
/	APPLICANT: Zhang, Jie	
/	APPLICANT: Qian, Xiaohong B.	
/	APPLICANT: Driemacq, Radoje T.	
/	TITLE OF INVENTION: No. US20020197679A1e1 Nucleic Acids and	
/	FILE REFERENCE: Polypeptides	
/	CURRENT APPLICATION NUMBER: US/10/098, 841	
/	PRIOR FILING DATE: 2002-03-13	
/	PRIOR APPLICATION NUMBER: 09/598, 042	
/	PRIOR FILING DATE: 2000-06-20	
/	PRIOR APPLICATION NUMBER: 09/552, 317	
/	PRIOR FILING DATE: 2000-04-25	
/	PRIOR APPLICATION NUMBER: 09/488, 725	
/	PRIOR FILING DATE: 2000-01-21	
/	NUMBER OF SEQ ID NOS: 331	
/	SOFTWARE: pc_FL_genes Version 1.0	
/	SEQ ID NO 271	
/	LENGTH: 2092	
/	TYPE: DNA	
/	ORGANISM: Homo sapiens	
/	FEATURES:	
/	NAME/KEY: CDS	
/	LOCATION: (139)..(1215)	
/	US-10-098-841-271	
Query Match	96.4%; Score 2039.4; DB 14; Length 2092;	
Best Local Similarity	99.2%; Pred. No. 0;	
Matches 2060; Conservative	0; Mismatches 16; Indels 1; Gaps 1	
OY	39 CAGGGCTGGAAGTGGGTGGATCCCGAGCTGCAGCAGACGCA-CGGGCCGGCCACTT	97
Dd	16 CCGGATCACCAAGCGTCATCCCGAGCTCGGACAGCAGCGCGGGCCGACCCT	75
OY	98 GCTGGTCCCCTGAGAGGCTCTGAGCCCCGGGGGGGCCCAGCGAAAGCGGGC	157
Dd	76 GCTGTGCTCCGAGAGGCTCTGAGCCCCGGGGGGGCCCAGCGAAAGCGGGC	135
OY	158 GAGATCGAGCACCCCTCTGCTGCTCTTGCGGGTTCCCTGTCAGAGAAGCGATTG	217
Dd	136 GAGATGGAAGCACCCCTCTGCTGCTCTTGCGGGTTCCCTGTCAGAGAAGCGGGTTG	195
OY	218 GAGTTGATGACAATAAGATACCGAGAGTCCTCCAGAAAAGAGTCCGAATGGGCC	277
Dd	196 GAGTTGATGACAATAAGATACCGAGAGTCCTCCAGAAAAGAGTCCGAATGGGCC	255
OY	278 CAGCCAGACTGCCCTGCTGTGTCCTCCAGAACCACTACTGCTCCAGATCGTGA	337
Dd	256 CAGCCAGACTGCCCTGCTGTGTCCTCCAGAACCACTACTGCTCCAGATCGTGA	315
OY	338 ACTGCTGTGGCACTGCTCCGCTCTTGAGGCCCTATATCTCTCTTGAGCCCGAGAGGC	397
Dd	316 ACTGCTGTGGCACTGCTCCGCTCTTGAGGCCCTATATCTCTCTTGAGCCCGAGAGGC	375

QY 398 GGGGGGCTTACAGAGCCCTTCCACTGCTCCCTTACAGAGCATAGATATACCTGCAAGGTGTAC 457
DB 376 GGGGGGGCTTACAGAGCCCTTCCACTGCTCCCTTACAGAGCATAGATATACCTGCAAGGTGTAC 435
QY 458 CCGGTTCAGAGAGCCCTGCGGTGTGTGAGCCCTTACAGAGCCCTGCGGTGTGTGAGCCCTTACAGAG 517
DB 436 CCGGTTCAGAGAGCCCTGCGGTGTGTGAGCCCTTACAGAGCCCTGCGGTGTGTGAGCCCTTACAGAG 495
QY 518 GTGGGTGCGCCCTGAGGTCTGTGCGGTGTGTGAGCCCTTACAGAGCCCTTACAGAGCCCTTACAGAG 577
DB 496 GTGGGTGCGCCCTGAGGTCTGTGCGGTGTGTGAGCCCTTACAGAGCCCTTACAGAGCCCTTACAGAG 555
QY 578 ACCCATGAGGAGCATATGACAGAGCCCTGAGTGCAGAAAGCGCCACCGATATCCCTGAGCCCTGAGGCT 637
DB 556 ACCCATGAGGAGCATATGACAGAGCCCTGAGTGCAGAAAGCGCCACCGATATCCCTGAGCCCTGAGGCT 615
QY 638 GCGGTGCTCTTCCGCGCAAGTGGCCACCGCCCTGCGGTGTGTGAGCCCTTACAGAGCCCTTACAGAG 697
DB 616 GCGGTGCTCTTCCGCGCAAGTGGCCACCGCCCTGCGGTGTGTGAGCCCTTACAGAGCCCTTACAGAG 675
QY 698 CTGCGTGTATCTCAAGCTGT 757
DB 676 CTGCGTGTATCTCAAGCTGT 735
QY 758 CTGAGAGAACTTGAAGAGACTCTCTGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 817
DB 736 CTGAGAGAACTTGAAGAGACTCTCTGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 795
QY 818 CAGCGGTGCTCAAGCTGT 877
DB 796 CAGCGGTGCTCAAGCTGT 855
QY 878 AAGGAGAGCCGATGT 937
DB 856 AAGGAGAGCCGATGT 915
QY 938 CCGCTTCAAGAGCTCGAGAGCTGT 997
DB 916 CCGCTTCAAGAGCTCGAGAGCTGT 975
QY 998 TTGCGCTGAGAGCTCTGCGCCCTTCCCTGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1057
DB 976 TTGCGCTGAGAGCTCTGCGCCCTTCCCTGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1035
QY 1058 CAGAGCTGAGAGCTCTGAGAGCTGT 1117
DB 1036 CAGAGCTGAGAGCTCTGAGAGCTGT 1095
QY 1118 ATGCGCTTGAAGCCCAAGCCGATCCCATCTCTGAGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGT 1177
DB 1096 ATGCGCTTGAAGCCCAAGCCGATCCCATCTCTGAGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGT 1155
QY 1178 CTGGGCTGTGAGAGAGCCGAG 1237
DB 1156 CTGGGCTGTGAGAGAGCCGAG 1215
QY 1238 GAGCAGCCCTTACAGAGCTGT 1297
DB 1216 GAGCAGCCCTTACAGAGCTGT 1275
QY 1298 CTCTCTGCTGT 1357
DB 1276 CTCTCTGCTGT 1335
QY 1358 AGCCTGT 1417
DB 1336 AGCCTGT 1395
QY 1418 TTGGGCTGT 1477
DB 1396 TTGGGCTGT 1455
QY 1478 GAGAGCAATATTTCCCTGTCTCAGAGATGACAAATCTGCGATCTTGTAGCTGACACACTT 1537

DB 1456 GAGAGCAATATTTCCCTGTCTCAGAGATGACAAATCTGCGATCTTGTAGCTGACACACTT 1515
QY 1538 TTCCATGACCATATAGTACATGTCTTACACTGAGGTAACCTTTGTACAGTGTGCGCTCCAC 1597
DB 1516 TTCCATGACCATATAGTACATGTCTTACACTGAGGTAACCTTTGTACAGTGTGCGCTCCAC 1575
QY 1598 TGATGT 1657
DB 1576 TGATGT 1635
QY 1658 CCTTTGATCTTGTATACCTTTTACAGAGAAAGGAGATCCCTGTGTGTGTGTGTGTGTGTGTGTGT 1717
DB 1636 CCTTTGATCTTGTATACCTTTTACAGAGAAAGGAGATCCCTGTGTGTGTGTGTGTGTGTGTGTGT 1695
QY 1718 CTCTCCCTGTGACAG 1777
DB 1696 CTCTCCCTGTGACAG 1755
QY 1778 TCTGTCTCTGTGTATAGAGATTTCTCTTCAAGGCTTACAGGCTTGTGTGTGTGTGTGTGTGTGTGT 1837
DB 1756 TCTGTCTCTGTGTATAGAGATTTCTCTTCAAGGCTTACAGGCTTGTGTGTGTGTGTGTGTGTGTGT 1815
QY 1838 TAAAGATCCAACTATAG 1897
DB 1816 TAAAGATCCAACTATAG 1875
QY 1898 CCAGGCTGTGTCAAGCATGAGGCTTGT 1957
DB 1876 CCAGGCTGTGTCAAGCATGAGGCTTGT 1935
QY 1958 CAGGCTGT 2017
DB 1936 CAGGCTGT 1995
QY 2018 GGTGTCAATATCCAGGCTCATCTCTAGGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2077
DB 1996 GGTGTCAATATCCAGGCTCATCTCTAGGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2055
QY 2078 GTGCTTAATTAAGAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2114
DB 2056 GTGCTTAATTAAGAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2092

RESULT 6
US-09-925-301-260
; Sequence 260, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: P106
; CURRENT APPLICATION NUMBER: US/09/925,301
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 260
; LENGTH: 2048
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (66)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (67)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-260

Query Match 83.3%; Score 1762.4; DB 9; Length 2048;
 Best Local Similarity 95.3%; Pred. No. 0;
 Matches 1876; Conservative 4; Mismatches 0; Indels 76; Gaps 4;

QY 161 ATGGAGCCACCCCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 220
 DB 156 ATGGAGCCACCCCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 215
 QY 221 TTGGATGACAACTTATAGTACCGAGCGTCCCGTCCAGAAAGCGGTTGAG 280
 DB 216 TTGGATGACAACTTATAGTACCGAGCGTCCCGTCCAGAAAGCGGTTGAG 275
 QY 281 CCCGACTGCCCCCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 340
 DB 276 CCCGACTGCCCCCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 304
 QY 341 GCTGTGCGCACTGCTCCCGTCTTGGAGCCCTATGCTCTCTGAGAGCCCGAGAGGCGG 400
 DB 305 -----GAGCCCGAGAGGCGG 322
 QY 401 CGGAGCTTACGAGGCTTGAAGCTTGAAGCAAGCACTGATATACCTGAGGTTACCC 460
 DB 323 CGGAGCTTACGAGGCTTGAAGCTTGAAGCAAGCACTGATATACCTGAGGTTACCC 380
 QY 461 GTCCAGAAAGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 520
 DB 381 GTCCAGAAAGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 440
 QY 521 GCTCGGCGCACTGAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 580
 DB 441 GCTCGGCGCACTGAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 500
 QY 581 CATGGGCAATGACAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 640
 DB 501 CATGGGCAATGACAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 559
 QY 641 GTCTCTTCCGAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 700
 DB 560 GTCTCTTCCGAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 619
 QY 701 CGTATCTCAAGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 760
 DB 620 CGTATCTCAAGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 679
 QY 761 GAGAACCTGAGAGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 820
 DB 680 GAGAACCTGAGAGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 739
 QY 821 GGTGCTCCGAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 880
 DB 740 GGTGCTCCGAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 799
 QY 881 GAGAGCATGCTGAGAGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 940
 DB 800 GAGAGCATGCTGAGAGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 859
 QY 941 TTTCAGAGCTGAGAGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1000
 DB 860 TTTCAGAGCTGAGAGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 919
 QY 1001 CCGTGAAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1060
 DB 920 CCGTGAAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 979
 QY 1061 GGTGAAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1120
 DB 980 GGTGAAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1039
 QY 1121 CCTTGAAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1180
 DB 1040 CCTTGAAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1099
 QY 1181 GGGCTGAAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1240

DB 1100 GGGCTGAAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1159
 QY 1241 CACCTTACTACAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1300
 DB 1160 CACCTTACTACAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1219
 QY 1301 CTCTGCTTGAAGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1360
 DB 1220 CTCTGCTTGAAGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1279
 QY 1361 CTGCTGAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1420
 DB 1280 CTGCTGAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1339
 QY 1421 GGTGCTTATGAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1480
 DB 1340 GGTGCTTATGAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1399
 QY 1481 GAGCAATATCCGCTTGAAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1540
 DB 1400 GAGCAATATCCGCTTGAAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1459
 QY 1541 CATGACATAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1600
 DB 1460 CATGACATAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1519
 QY 1601 TGCTGCTTGAAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1660
 DB 1520 TGCTGCTTGAAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1579
 QY 1661 TTGTATCTTGAAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1720
 DB 1580 TTGTATCTTGAAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1639
 QY 1721 TCCCTGCAAGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1780
 DB 1640 TCCCTGCAAGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1699
 QY 1781 GTCTCTTATTAAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1840
 DB 1700 GTCTCTTATTAAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1759
 QY 1841 GAATCCAACTAAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1900
 DB 1760 GAATCCAACTAAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1819
 QY 1901 GGCCTGCAAGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1960
 DB 1820 GGCCTGCAAGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1879
 QY 1961 GGTCTGCTTGAAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 2020
 DB 1880 GGTCTGCTTGAAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1939
 QY 2021 CCAGATCCAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 2080
 DB 1940 CCAGATCCAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 1999
 QY 2081 CCTAATAAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 2140
 DB 2000 CCTAATAAGGCTTGGCTGCTCTTCCGAGGTTCCCTGTCCAGAGAAAGCGGTTGAG 2035

RESULT 7
 US-10-024-828-3
 ; Sequence 3, Application US/10024828
 ; Publication No. US20030036051A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Virex, Duke
 ; APPLICANT: Bird, Timothy A.
 ; APPLICANT: Andersen, Dirk M.
 ; APPLICANT: Marken, John S.


```

Db      301 GTCCAGGAGGCGCTGCGCTGCTGAGAGCCCTACGCGCGGCTGCCCGGCAAAACATGTG 360
Qy      521 GCTGGGCGCCACTGAGGCTCCTGAGCTGTAACCAAGTCTCTTACGCTTTTTCATCTGGACC 580
Db      361 GCTGGGCGCCACTGAGGCTCCTGAGCTGTAACCAAGTCTCTTACGCTTTTTCATCTGGACC 420
Qy      581 CATGGGGAATGCAACAGCTGCTGAGCAAGCGGCGCAAGCTTACCTGAGCTGAGGCTGCC 640
Db      421 CATGGGGAATGCAACAGCTGCTGAGCAAGCGGCGCAAGCTTACCTGAGCTGAGGCTGCC 480
Qy      641 GTGCTCTTCCGCGCAAGAGGCGCAAGCGGCGCAAGCTTACCTGAGCTGAGGCTGCC 700
Db      481 GTGCTCTTCCGCGCAAGAGGCGCAAGCGGCGCAAGCTTACCTGAGCTGAGGCTGCC 540
Qy      701 CGTGAATCTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 760
Db      541 CGTGAATCTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Qy      761 GAGAACTGAGAGGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 820
Db      601 GAGAACTGAGAGGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Qy      821 GCGTGGCCCACTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 880
Db      661 GCGTGGCCCACTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Qy      881 GCAGCCGATGTCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 940
Db      721 GCAGCCGATGTCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Qy      941 TTCAGAGACTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1000
Db      781 TTCAGAGACTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Qy      1001 CCTGGAGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGG 1060
Db      841 CCTGGAGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGG 900
Qy      1061 GGTGAAGGCTCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1120
Db      901 GGTGAAGGCTCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Qy      1121 CCTTGAAGGCTCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1180
Db      961 CCTTGAAGGCTCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Qy      1181 GGGCTGGAAGAGGCGGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1234
Db      1021 GGGCTGGAAGAGGCGGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1074

RESULT 9
US-10-649-156-9
; Sequence 9, Application US/10649156
; Publication No. US20040038346A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Williamson, Mark
; TITLE OF INVENTION: No. US20040038346A1el Human Protein Kinases and Uses
; FILE REFERENCE: 35800/209996
; CURRENT APPLICATION NUMBER: US/10/649,156
; PRIOR FILING DATE: 2003-08-27
; PRIOR APPLICATION NUMBER: US/09/799,875
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: 60/182,059
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/659,287
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 9
; LENGTH: 1074
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-649-156-9

Query Match      50.5%; Score 1069.2; DB 13; Length 1074;
Best Local Similarity 99.7%; Pred. No. 6,6e-308;
Matches 1071; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      161 ATGAGAGCCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 220
Db      1 ATGAGAGCCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
Qy      221 TTGAGTGAACACTTAATGATACGAGCGTCCGCTCCAGAAAGAGCTGAGTGGAGCCAG 280
Db      61 TTGAGTGAACACTTAATGATACGAGCGTCCGCTCCAGAAAGAGCTGAGTGGAGCCAG 120
Qy      281 CCCAGACTGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 340
Db      121 CCCAGACTGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Qy      341 GCTGTGGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 400
Db      181 GCTGTGGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Qy      401 CGGAGCTTACAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 460
Db      241 CGGAGCTTACAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Qy      461 GTTCAGAGAGCCCTGAGCGGCTGAGAGCCCTGAGAGCGGCGGCTGCGGCGGCGGCAAGAGCTG 520
Db      301 GTTCAGAGAGCCCTGAGCGGCTGAGAGCCCTGAGAGCGGCGGCTGCGGCGGCGGCAAGAGCTG 360
Qy      521 GCTGGGCGCCACTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 580
Db      361 GCTGGGCGCCACTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Qy      581 CATGGGGAATGCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 640
Db      421 CATGGGGAATGCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Qy      641 GTGCTCTTCCGCGCAAGAGGCGCAAGCGGCGCAAGCGGCGCAAGCGGCGCAAGCGGCGCAAG 700
Db      481 GTGCTCTTCCGCGCAAGAGGCGCAAGCGGCGCAAGCGGCGCAAGCGGCGCAAGCGGCGCAAG 540
Qy      701 GGTGAATCTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 760
Db      541 GGTGAATCTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Qy      761 GAGAACTGAGAGGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 820
Db      601 GAGAACTGAGAGGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Qy      821 GCGTGGCCCACTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 880
Db      661 GCGTGGCCCACTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Qy      881 GCAGCCGATGTCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 940
Db      721 GCAGCCGATGTCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Qy      941 TTCAGAGACTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1000
Db      781 TTCAGAGACTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Qy      1001 CCTGAGAGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGG 1060
Db      841 CCTGAGAGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGG 900
Qy      1061 GGTGAAGGCTCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1120
Db      901 GGTGAAGGCTCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960

```


PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 525
LENGTH: 562
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (515)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (526)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (557)
OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-525

Query Match 17.5%; Score 369.4; DB 9; Length 562;
Best Local Similarity 97.2%; Pred. No. 2,4e-99;
Matches 378; Conservative 6; Mismatches 4; Indels 1; Gaps 1;

QY 1 GGAAGCGGCTCCGCGCGGCTCCGCTGCTAGAGACCGGCGAGGCTGAGCTGGCTGGGA 60
DB 161 GGAAGCGGCTCCGCGCGGCTCCGCTGCTAGAGACCGGCGAGGCTGAGCTGGCTGGGA 220
QY 61 TCCGAGCTCGGAG 119
DB 221 TCCGAGCTCGGAG 280
QY 120 GCCCG 179
DB 281 GCCCG 340
QY 180 CTCTCTCTGCGGCTTCT 239
DB 341 CTCTCTCTGCGGCTTCT 400
QY 240 CCGAGCGGCTCCGCGCGGAG 239
DB 401 CCGAGCGGCTCCGCGCGGAG 460
QY 300 TGTTCCT 359
DB 461 TGTTCCT 520
QY 360 GTCTTGGGCGGCTTGTCT 388
DB 521 GTTTCGCGGCTTGTCT 549

RESULT 13

US-09-803-719-519.
Sequence 519, Application US/09803719
Publication No. US20030044783A1
GENERAL INFORMATION:
APPLICANT: Williams, Jaime T.
APPLICANT: Escobedo, Lewis T.
APPLICANT: Innis, Michael A.
APPLICANT: Garcia, Pablo Dominguez
APPLICANT: Sudduth-Klinger, Julie
APPLICANT: Reinhard, Christoph
APPLICANT: Giese, Klaus
APPLICANT: Randazzo, Filippo
APPLICANT: Kennedy, Giulia C.
APPLICANT: Poc, David
APPLICANT: Kassam, Altaf
APPLICANT: Lamson, George
APPLICANT: Drmanac, Radoje
APPLICANT: Cirkenjakov, Radomir
APPLICANT: Dickson, Mark

APPLICANT: Drmanac, Snezana
APPLICANT: Labat, Ivan
APPLICANT: Leshkowitz, Dena
APPLICANT: Kita, David
APPLICANT: Garcia, Veronica
APPLICANT: Jones, Lee William
APPLICANT: Stache-Crain, Birgit
TITLE OF INVENTION: Human Genes and Gene Products
FILE REFERENCE: 1624.002
CURRENT APPLICATION NUMBER: US/09/803,719
CURRENT FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 60/188,609
PRIOR FILING DATE: 2000-03-09
NUMBER OF SEQ ID NOS: 2396
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 519
LENGTH: 396
TYPE: DNA
ORGANISM: Homo sapiens
US-09-803-719-519

Query Match 16.7%; Score 354.2; DB 10; Length 396;
Best Local Similarity 96.4%; Pred. No. 7.2e-95;
Matches 373; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

QY 70 CCGCAGAGAGGCA-CGGGCGGCGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 128
DB 9 CCGCAGAGAGGCACTTGGCTTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 68
QY 129 GCGCGCGGCG 188
DB 69 GCGCGCGGCG 128
QY 189 CGGCTTCCT 248
DB 129 CGGCTTCCT 188
QY 249 CCGTCCAGAAAGAGCTGCAAGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 308
DB 189 CCGTCCAGAAAGAGCTGCAAGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 248
QY 309 TGAGCCCACT 368
DB 249 TGAGCCCACT 308
QY 369 CCTATGCT 428
DB 309 CCTATGCT 368
QY 429 CAGGCACTGAGTACTCTGCAAGGTGT 455
DB 369 CAGGCACTGAGTACTCTGCAAGGTGT 395

RESULT 14

US-09-873-367C-341/C
Sequence 341, Application US/09873367C
Publication No. US20030165839A1
GENERAL INFORMATION:
APPLICANT: Young, Paul
APPLICANT: Soppet, Daniel
APPLICANT: Andress, Gregory
APPLICANT: Augustus, Meena
APPLICANT: Ebner, Reinhard
APPLICANT: Carter, Kenneth
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
FILE REFERENCE: 689290-64
CURRENT APPLICATION NUMBER: US/09/873,367C
CURRENT FILING DATE: 2003-04-29
PRIOR APPLICATION NUMBER: U.S. 60/236,891
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: U.S. 60/236,842

PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: U.S. 60/244,867
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: U.S. 60/245,084
PRIOR FILING DATE: 2000-11-01
NUMBER OF SEQ ID NOS: 1067
SOFTWARE: PatentIn version 3.0
SEQ ID NO 341
LENGTH: 353
TYPE: DNA
ORGANISM: Homo sapiens
US-09-873-367C-341

Query Match 14.9%; Score 315.4; DB 10; Length 353;
Best Local Similarity 95.2%; Pred. No. 2.5e-83;
Matches 336; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

QY 1756 TGGAACTGTGTT-CCGACATCTCTGCTCTTGAATTAAGAGATTCTCTCCAGGCTT 1814
Db 353 TGGAACTGTGTTCCCGACATCTCTGCTCTTGAATTAAGAGATTCTCTCCAGGCTT 294
QY 1815 AAGCTGGATTTGGGCGCAGATTAAGATTCGAATCTAGAGCTAGTCTTGTCTAACT 1874
Db 293 AAGCTGGATTTGGGCGCAGATTAAGATTCGAATCTAGAGCTAGTCTTGTCTAACT 234
QY 1875 CAAGACTGTTCTGAATGAGGCTCCAGGCTGTCACATGAGGCTTCTGACCTGAGCAG 1934
Db 233 CAAGACTGTTCTGAATGAGGCGCAGGCTGTCACATGAGGCTTCTGACCTGAGCAG 174
QY 1935 CAAGTTGAGGAGCAGAGATTAGGCGAGGCTGCTGCTGTCGACCTGAGAAAGTCCAGGT 1994
Db 173 CAAGTTGAGGAGCAGAGATTAGGCGAGGCTGCTGCTGTCGACCTGAGAAAGTCCAGGT 114
QY 1995 GGGACCTCTTCTGGGAGCACTTGGGCTCCAAATCCAGGCTCACTCTAGGTTTGGAT 2054
Db 113 GGGACCTCTTCTGGGAGCACTTGGGCTCCAAATCCAGGCTCACTCTAGGTTTGGAT 54
QY 2055 ACCATGATGATGATGTTTACCTGCTTAATTAAGAGAAATTAATAA 2107
Db 53 ACCATGATGATGATGTTTACCTGCTTAATTAAGAGAAATTAATAA 1

RESULT 15
US-10-221-278-116
Sequence 116, Application US/10221278
Publication No. US20040034208A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: No. US20040034208A1 Nucleic Acids and Polypeptides
FILE REFERENCE: 21272-045
CURRENT APPLICATION NUMBER: US/10/221,278
CURRENT FILING DATE: 2002-09-06
PRIOR APPLICATION NUMBER: 09/693,267
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 09/665,363
PRIOR FILING DATE: 2000-09-19
PRIOR APPLICATION NUMBER: 09/616,847
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 09/596,193
PRIOR FILING DATE: 2000-06-17
PRIOR APPLICATION NUMBER: 09/574,454
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: 09/519,705
PRIOR FILING DATE: 2000-03-07
NUMBER OF SEQ ID NOS: 752
SEQ ID NO 116
LENGTH: 1909
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (135)..(941)
US-10-221-278-116

Query Match 12.2%; Score 258.4; DB 13; Length 1909;
Best Local Similarity 60.3%; Pred. No. 4.8e-66;
Matches 444; Conservative 0; Mismatches 291; Indels 1; Gaps 1;

QY 370 CTATGCTCTCTGAGAGCCGAGAGGCGGCGGCTTACACAGGCTGCACTGCTTAC 429
Db 103 CTACCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCT 162
QY 430 AGGCACTGATATAC-CTGCAAGGTGTACCCCGTCAGAGAACCTGCGCTGAGAC 488
Db 163 TGGAGCGCGAGGTGCGGCAAGGTGTGTCCATTAACACTACAGGACAAATCAGGCT 222
QY 489 CTTACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 548
Db 223 CTTACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 282
QY 549 CCGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 608
Db 283 AACCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 342
QY 609 CCGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 668
Db 343 GCGGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 402
QY 669 TGGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 728
Db 403 TGGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 462
QY 729 TGGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 788
Db 463 TGGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 522
QY 789 CTGAGCAGATGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 848
Db 523 AGGCGGAAATGATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 582
QY 849 TACTGACTGAGGCTCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 908
Db 583 TCTTCAACACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 642
QY 909 CGCTCTTCAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 968
Db 643 TGTCTTCAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 702
QY 969 TCGGCAAGATCCGCGCGGCGGCTTACGCTTGTGCTGCTGCTGCTGCTGCTGCTGCT 1028
Db 703 TCTTCAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 762
QY 1029 GTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1088
Db 763 GCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 822
QY 1089 TCTTCAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1104
Db 823 TACTGACCTCTGCTT 838

Search completed: August 29, 2004, 14:34:59
Job time : 972 secs


```
Db      121  CCGAAGTACGCCCCCTGCTGCTGTTGCCCCCTGAGCCCACTACTGCTCCAGATGTCGCACT 180
Qy      341  GGTGAGGCACTGCTCCGCTGTTGAGGCTTATGTCCTCTGAGAGCCGAGAGAGGCGG 400
Db      181  GGTGAGGCACTGCTCCGCTGTTGAGGCTTATGTCCTCTGAGAGCCGAGAGAGGCGG 240
Qy      401  CCGGCTTACAGAGCCCTGCTGAGCCCTTACAGAGCACTGATATACCTGCAAGGTGATCCC 460
Db      241  CCGGCTTACAGAGCCCTGCTGAGCCCTTACAGAGCACTGATATACCTGCAAGGTGATCCC 300
Qy      461  GTCCAGAGAGCCCTGAGGCTGCTGAGAGCCCTTACAGAGCACTGATATACCTGCAAGGTG 520
Db      301  GTGACAGAGAGCCCTGAGGCTGCTGAGAGCCCTTACAGAGCACTGATATACCTGCAAGGTG 360
Qy      521  GTTCCGAGAGAGCCCTGAGGCTGCTGAGAGCCCTTACAGAGCACTGATATACCTGCAAGGTG 580
Db      361  GCTCCGAGAGAGCCCTGAGGCTGCTGAGAGCCCTTACAGAGCACTGATATACCTGCAAGGTG 420
Qy      581  CATGGGAGAGAGCCCTGAGGCTGCTGAGAGCCCTTACAGAGCACTGATATACCTGCAAGGTG 640
Db      421  CCGGAGAGAGAGCCCTGAGGCTGCTGAGAGCCCTTACAGAGCACTGATATACCTGCAAGGTG 480
Qy      641  GTGCTCTTCCGCAAGAGAGCCCTGAGGCTGCTGAGAGCCCTTACAGAGCACTGATATACCTG 700
Db      481  GCGCTCTTCCGCAAGAGAGCCCTGAGGCTGCTGAGAGCCCTTACAGAGCACTGATATACCTG 540
Qy      701  CGTGAATCTAGAGAGAGCCCTGAGGCTGCTGAGAGCCCTTACAGAGCACTGATATACCTG 760
Db      541  CGTGAATCTAGAGAGAGCCCTGAGGCTGCTGAGAGCCCTTACAGAGCACTGATATACCTG 600
Qy      761  GAGAACTGAGAGAGCCCTGAGGCTGCTGAGAGCCCTTACAGAGCACTGATATACCTG 820
Db      601  GAGAACTGAGAGAGCCCTGAGGCTGCTGAGAGCCCTTACAGAGCACTGATATACCTG 660
Qy      821  GCGTCCGCAAGAGAGCCCTGAGGCTGCTGAGAGCCCTTACAGAGCACTGATATACCTG 880
Db      661  GCGTCCGCAAGAGAGCCCTGAGGCTGCTGAGAGCCCTTACAGAGCACTGATATACCTG 720
Qy      881  GCGTCCGCAAGAGAGCCCTGAGGCTGCTGAGAGCCCTTACAGAGCACTGATATACCTG 940
Db      721  GCGTCCGCAAGAGAGCCCTGAGGCTGCTGAGAGCCCTTACAGAGCACTGATATACCTG 780
Qy      941  TTCAGAGAGAGAGCCCTGAGGCTGCTGAGAGCCCTTACAGAGCACTGATATACCTG 1000
Db      781  TTCAGAGAGAGAGCCCTGAGGCTGCTGAGAGCCCTTACAGAGCACTGATATACCTG 840
Qy      1001  CCGTCCGCAAGAGAGCCCTGAGGCTGCTGAGAGCCCTTACAGAGCACTGATATACCTG 1060
Db      841  CCGTCCGCAAGAGAGCCCTGAGGCTGCTGAGAGCCCTTACAGAGCACTGATATACCTG 900
Qy      1061  GCGTCCGCAAGAGAGCCCTGAGGCTGCTGAGAGCCCTTACAGAGCACTGATATACCTG 1120
Db      901  GCGTCCGCAAGAGAGCCCTGAGGCTGCTGAGAGCCCTTACAGAGCACTGATATACCTG 960
Qy      1121  CCGTCCGCAAGAGAGCCCTGAGGCTGCTGAGAGCCCTTACAGAGCACTGATATACCTG 1180
Db      961  CCGTCCGCAAGAGAGCCCTGAGGCTGCTGAGAGCCCTTACAGAGCACTGATATACCTG 1020
Qy      1181  GCGTCCGCAAGAGAGCCCTGAGGCTGCTGAGAGCCCTTACAGAGCACTGATATACCTG 1234
Db      1021  GCGTCCGCAAGAGAGCCCTGAGGCTGCTGAGAGCCCTTACAGAGCACTGATATACCTG 1074
```

```
RESULT 3
AL578892/c 1201 bp mRNA linear EST 01-JUN-2003
LOCUS AL578892 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
DEFINITION cDNA clone CS0DK005YC05 3'-PRIME, mRNA sequence.
ACCESSION AL578892
VERSION AL578892.2 GI:31317080
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jeessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 16, 2001 this sequence version replaced gi:12943405.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10000.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DK005AB03NP1;cluster=10000.f. Contact :
Peng Liang Email : fliang@life.techn.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DK005AB03NP1.
Location/Qualifiers

FEATURES

source
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK005YC05"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/clone_1ib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 44.5%; Score 940.6; DB 9; Length 1201;
Best Local Similarity 96.1%; Pred. No. 8e-169;
Matches 943; Conservative 24; Mismatches 13; Indels 1; Gaps 1;
Qy 1096 CCCCTGGCTGCGACAGAGCCGATGCTTACGCCCCAAGCCGATCCCATCTTGGAGAGC 1155
Db 992 MCCCTGSGCGGMMAGACCCRAKSCCTTDSFCCACCCCTTCCCTTC-KGAGAGC 934
Qy 1156 TGCCCAAGTGGTCCCGATGAGAGCTGGGCTGAGAGAGAGAGAGAGAGAGAGAG 1215
Db 933 TGCCVAGGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 874
Qy 1216 AGAAGTGGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1275
Db 873 AGAAGTGGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 814
Qy 1276 GAGTTTGGGAGTACCTCAAGCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1335
Db 813 GAGTTTGGGAGTACCTCAAGCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 754
Qy 1336 TTCCAG 1395
Db 753 TTCCAG 694
Qy 1396 TTCCAG 1455
Db 693 TTCCAG 634
Qy 1456 TGAGAGTACAG 1515
Db 633 TGAGAGTACAG 574
Qy 1516 CATCTTGGAGTACAG 1575
Db 573 CATCTTGGAGTACAG 514
Qy 1576 TTGTACAGAGTCCGAGTCCAG 1635
Db 513 TTGTACAGAGTCCGAGTCCAG 454

[illegible]

FEATURES

Source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS00DL0023707"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dAT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

Query Match	42.3%;	Score 895.2;	DB 13;	Length 1075;
Best Local Similarity	-95.7%;	Pred. No. 3.6e-160;		
Matches 929;	Conservative 10;	Mismatches 30;	Indels 2;	Gaps 2;

```

QY      1126 AGCCCAACCCGATCCCATCTCTGGAGAGCGCCGACAGTGGTCCCTATGATGATCGGGGCT 1185
Db      970 ATCCCTTASTTAACCCCMCCCATTTTGGAGCTCCCACTGTGCTCCATATGACTGGGGGT 911
QY      1186 GGAGGAAGCCAGGAGAGGAGGAGACAGAGAGTGTTCTGTATGCTTAGACACACC 1245
Db      910 KGAAGAAACCAAGGAAAGGA -GGAAPACAGAGAAKTGTTCTKATGCTARACACACC 852
QY      1246 TACTACACGCTCAGCTGCAACAGTGGATTGATTTGGGGGTAACTCAAGCCTTCCT 1305
Db      851 TACTACACGCTCAGCTGCAACAGTGGATTGATTTGGGGGTAACTCAAGCCTTCCT 792
QY      1306 GCGCTGTAACTGAGGCAAACTTCAAGTCTTCCAGAAAGGAGAAAGGCAAAAGCCTGTG 1365
Db      791 GCGCTGTAACTGAGGCAAACTTCAAGTCTTCCAGAAAGGAGAAAGGCAAAAGCCTGTG 732
QY      1366 TGAAGTGTGCTGTATACATCTGCTTGTGTTCCACACATGACAGTTCTCTGGGTGC 1425
Db      731 TGAAGTGTGCTGTATACATCTGCTTGTGTTCCACACATGACAGTTCTCTGGGTGC 672
QY      1426 TTATCAGGTGCCAACCCTCTTCTTCGTGGTCTGGGAGTACAGAGTGAGCAAGAGACAA 1485
Db      671 TTATCAGGTGCCAACCCTCTTCTTCGTGGTCTGGGAGTACAGAGTGAGCAAGAGACAA 612
QY      1486 TATTCCTGCTCACAAGAGTGAACAATGCGATCCTTGAAGTGACAAACCTTTTCATGA 1545

```

Db	611	TATTCCTGCTCAGACAGATGACMACTGGACATCCTTAGAGCTGACAAACCTTTTCATGA	552
OY	1546	CCATAGGTCACTGTCTTCACTACGTGGGTAACATTGTACACAGTGTGCGGCTTCACATGATGCG	1605
Db	551	CCATAGGTCACTGTCTTCACTACGTGGGTAACATTGTACACAGTGTGCGGCTTCACATGATGCG	492
OY	1606	GTCGTCAAGGACCTGTCTGTCCAAAGAACATGCCCTTTCACAAACAAACACAGTGCCTTTGTA	1665
Db	491	GTCGTCAAGGACCTGTCTGTCCAAAGAACATGCCCTTTCACAAACAAACACAGTGCCTTTGTA	432
OY	1666	TCTGTACCTTTTCAGAAAGGAGATGCCGTGTGCCAAAGCTTCAGGCTCTTCCCC	1725
Db	431	TCTGTACCTTTTCAGAAAGGAGATGCCGTGTGCCAAAGCTTCAGGCTCTTCCCC	372
OY	1726	TGCAACTCAGAGCCCAAGCCCACTCACTCTGGGAACTGTGTTCCAGACATCTCTGCT	1785
Db	371	TGCAACTCAGAGCCCAAGCCCACTCACTCTGGGAACTGTGTTCCAGACATCTCTGCT	312
OY	1786	CTTATATTAAGATTTTCCTCTCAGAGCTTAAGCTCGGATTTTGGGCGAGATTAAGATC	1845
Db	311	CTTATATTAAGATTTTCCTCTCAGAGCTTAAGCTCGGATTTTGGGCGAGATTAAGATC	252
OY	1846	CAAACTATGAGGCTAGTCTTGTCTAACTCAAGACTGTCTGATGAGGTCAGAGCT	1905
Db	251	CAAACTATGAGGCTAGTCTTGTCTAACTCAAGACTGTCTGATGAGGTCAGAGCT	192
OY	1906	GTCACACCTTGGGGCTTTCGACCTAGACCAAGGTTGAGGACAGATGAGGAGGCT	1965
Db	191	GTCACACCTTGGGGCTTTCGACCTAGACCAAGGTTGAGGACAGATGAGGAGGCT	132
OY	1966	GTCCTGTAGGACCTGGAAGTCCAGAGTGGAGCTTTCTGGGACACTTGGGGTCCACA	2025
Db	131	GTCCTGTAGGACCTGGAAGTCCAGAGTGGAGCTTTCTGGGACACTTGGGGTCCACA	72
OY	2026	ATCCAGGTCATCTCTAGGTTTGGATTACC-ATGATATATGATGTTTACCTGTGCCCTA	2084
Db	71	ATCCAGGTCATCTCTAGGTTTGGATTACCMATGATATGATGTTTACCTGTGTTA	12
OY	2085	ATAAAGAGAA 2095	
Db	11	ATAAAGAGAA 1	

RESULT 6

	EST 23-MAY-2003
AL525890/c	1051 bp mRNA linear
LOCUS AL525890	
AL525890 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens	
DEFINITION	
CDNA clone CSDDC013YJ08 3-PRIME, mRNA sequence.	

ACCESSION	AL525890	GI:31063754
VERSION	AL525890.2	
KEYWORDS	EST	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1. (bases 1 to 1051)	
AUTHORS	I. W.B., Gruber,C., Jesse,J. and Polayes,D.	
TITLE	Full-length cdna libraries and normalization	
JOURNAL	unpublished (2001)	
COMMENT	On Feb 13, 2001 this sequence version replaced gi:12789383.	

BP 191 91006 Evr1 cecex - rance
Email : segregen@genomecope.cnrs.fr Web : www.genoscope.cnrs.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10000.f For
more information about this cluster, see
[http://www.genoscope.cnrs.fr/
cgi-bin/cluster.cgi?seq=CS00C013D04NP1&cluster=10000.f](http://www.genoscope.cnrs.fr/cgi-bin/cluster.cgi?seq=CS00C013D04NP1&cluster=10000.f) Contact :
Peng Liang Email : fliang@life technologies.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
Paradeway Avenue Genome sequence ID : CS00C013D04NP1.
Location/Qualifiers

FEATURES

source 1. 1051

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS00C013Y08"
/issue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

```

Query Match 42.2%; Score 893; DB 9; Length 1051;
 Best Local Similarity 95.7%; Pred. No. 9.3e-160;
 Matches 908; Conservative 25; Mismatches 14; Indels 2; Gaps 2;

```

QY 1122 CCTAGCCCCACCCGATCCCATCTCTGGAGAGGCTGCCAGGTGCTCCCTGATGACTGG 1181
DB CCTAGCTCCAMCCGATCCCATCTCTGGAGAGGCTGCCAGGTGCTCCCTGATGACTGG 890
QY 1132 GCTGAGGAGAGCCGAGGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1241
DB GCTGAGGAGAGCCGAGGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 830
QY 1242 ACCCTACTACAGCTCAGCTGCAACAGTGATGATGATGATGATGATGATGATGATG 1301
DB ACCCTACTACAGCTCAGCTGCAACAGTGATGATGATGATGATGATGATGATGATG 770
QY 1302 TCTGCTCTGAGCTGAGCCAAACCTTGAAGCTTCCAGAGAGAGAGAGAGAGAGAG 1361
DB TCTGCTCTGAGCTGAGCCAAACCTTGAAGCTTCCAGAGAGAGAGAGAGAGAGAG 710
QY 1362 TGTGTGAGTGTCTGTGTACACATCTGCTTGTTCACACATGAGCTGCTGCTGG 1421
DB TGTGTGAGTGTCTGTGTACACATCTGCTTGTTCACACATGAGCTGCTGCTGG 650
QY 1422 GTGCTTATCAGGTGCGCAAGCCTGCTTCTCGGTCTGGAGTACAGCAGTGAAGAG 1481
DB GTGCTTATCAGGTGCGCAAGCCTGCTTCTCGGTCTGGAGTACAGCAGTGAAGAG 590
QY 1482 ACAATATTCCTGCTCAGAGATGACAACTGGCATCTTGGAGTACAACTTTCC 1541
DB ACAATATTCCTGCTCAGAGATGACAACTGGCATCTTGGAGTACAACTTTCC 530
QY 1542 ATGACCATAGTCACTGTGTACACTGTGTACACTGTGTACACTGTGTACACTGT 1601
DB ATGACCATAGTCACTGTGTACACTGTGTACACTGTGTACACTGTGTACACTGT 470
QY 1602 GCTGTGCTAGGACCTCTGTCCAGAGCAATCCCTTTCAAAACCAACCACTGCTT 1661
DB GCTGTGCTAGGACCTCTGTCCAGAGCAATCCCTTTCAAAACCAACCACTGCTT 410
QY 1662 TGTATCTTGTACCTTTTCAGAGAAAGGAGTATCCCTGAGCAAGGCTCCAGCTCT 1721
DB TGTATCTTGTACCTTTTCAGAGAAAGGAGTATCCCTGAGCAAGGCTCCAGCTCT 350
QY 1722 CCCCTGCACTCAGAGCCCAAGCCAGCTCACTCTGGAAGTGTGTCCAGATCTCTG 1781
DB CCCCTGCACTCAGAGCCCAAGCCAGCTCACTCTGGAAGTGTGTCCAGATCTCTG 290
QY 1782 TCTCTTGAATTAAGAGATCTCTCTTCCAGGCTTAAGCTTGGGCTCAAGATTAAG 1841
DB TCTCTTGAATTAAGAGATCTCTCTTCCAGGCTTAAGCTTGGGCTCAAGATTAAG 230
QY 1842 AATCCAACTATGAGCTAGTCTTGTCTAATCAAGCTGTTCTGAGATGAGGCTCAG 1901
DB AATCCAACTATGAGCTAGTCTTGTCTAATCAAGCTGTTCTGAGATGAGGCTCAG 170
QY 1902 GCTGTCAACCATGAGGCTTCTGACCTGAGCAAGGTTGAGGAGACAGATTAGCAGG 1961
DB GCTGTCAACCATGAGGCTTCTGACCTGAGCAAGGTTGAGGAGACAGATTAGCAGG 110

```

QY 1962 GTCTGTCTGTGAGCCACTGAGAAAGTCCAGAGTGGACTCTTCTGAGGAGCACTTGGGCTC 2021
 DB GTCTGTCTGTGAGCCACTGAGAAAGTCCAGAGTGGACTCTTCTGAGGAGCACTTGGGCTC 51

QY 2022 CACAAATCCAGGTCCACTACTAGTTTGTGATACCATAGATGATGATG 2070
 DB BAAATSCVGGTCAATAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 3

RESULT 7

EX325344/c

LOCUS EX325344 947 bp mRNA linear EST 02-MAY-2003
 DEFINITION Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED

ACCESSION Homo sapiens cDNA clone CS00L004YF04 3-PRIME, mRNA sequence.
 VERSION EX325344.1 GI:3036457

KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 947)
 L.I.B., Gruber, C., Jesse, J. and Polayes, D.
 Full-length cDNA libraries and normalization

COMMENT Unpublished (2001)
 Contact: Genoscope

Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 10000.f For
 more information about this cluster, see

http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0AL004DC02NP1&cluster=10000.f. Contact :

Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0AL004DC02NP1.
 Location/Qualifiers

FEATURES

source

1. 947

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS00L004YF04"
/cell_type="B CELLS (RAMOS CELL LINE)"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

Query Match 42.2%; Score 892; DB 13; Length 947;
 Best Local Similarity 96.4%; Pred. No. 1.4e-159;
 Matches 907; Conservative 22; Mismatches 10; Indels 2; Gaps 2;

```

QY 1142 CATCTCTGGAGGCTGCGGAGTGTCTCTGATGAGCTGAGGAGGAGGAGGAG 1201
DB CATCTCTGGAGGCTGCGGAGTGTCTCTGATGAGCTGAGGAGGAGGAGGAGGAG 888
QY 947 CTTCTCTGGAGGCTGCGGAGTGTCTCTGATGAGCTGAGGAGGAGGAGGAG 888
DB CTTCTCTGGAGGCTGCGGAGTGTCTCTGATGAGCTGAGGAGGAGGAGGAGGAG 828
QY 1202 GAGGAGGAGGAGGAGGAGGAGTGTCTGATGAGGAGGAGGAGGAGGAGGAG 1261
DB GAGGAGGAGGAGGAGGAGGAGTGTCTGATGAGGAGGAGGAGGAGGAGGAGGAG 828
QY 887 RAGGAGGAGGAGGAGGAGGAGTGTCTGATGAGGAGGAGGAGGAGGAGGAG 828
DB RAGGAGGAGGAGGAGGAGGAGTGTCTGATGAGGAGGAGGAGGAGGAGGAGGAG 1321
QY 827 GCCAACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 768
DB GCCAACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 768
QY 1322 AAACCTTCACTGCTTCCAGAGAGGAGAGAGGAGGAGGAGGAGGAGGAGGAG 1381
DB AAACCTTCACTGCTTCCAGAGAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 708

```

QY	1382	CACATCTGCTTTGTTCCACACACATGAGAGTTCCGTGGGAGTCTTATAGGAGCCAC	1441
Db	707	CACATCTGCTTTGTTCCACACACAAATAGTTCTGCTGGGAGTCTTATAGGAGCCAC <td>648</td>	648
QY	1442	CCGTGTCGGTGTGGGAGTACAGCAGTAGAGCAAAAGACAAATATTCCTGCTCAG <td>1501</td>	1501
Db	647	CCGTGTCGGTGTGGGAGTACAGCAGTAGAGCAAAAGACAAATATTCCTGCTCAG <td>588</td>	588
QY	1502	AGATGACAAATGCGCATCTTTGAGCTGACAAACATTTCCATGACAAATAGTACATG <td>1561</td>	1561
Db	587	AGATGACAAATGCGCATCTTTGAGCTGACAAACATTTCCATGACAAATAGTACATG <td>528</td>	528
QY	1562	ACACTGGGTACATTTGTATCCAGTGTGGCTCCACTGATAGCTGTGCTCAGGACCT <td>1621</td>	1621
Db	527	ACACTGGGTACATTTGTATCCAGTGTGGCTCCACTGATAGCTGTGCTCAGGACCT <td>469</td>	469
QY	1622	GTCCAGGACAAATCCCTTTTCACAACAAACAGCTGCTTTGTATCTTGTACCTTTT <td>1681</td>	1681
Db	468	GTCCAGGACAAATCCCTTTTCACAACAAACAGCTGCTTTGTATCTTGTACCTTTT <td>409</td>	409
QY	1682	AGAAAGGAGGTATCCCTGTGACCAAGGCTTCAGGCTCTTCCCTGCACTCAGGACCA <td>1741</td>	1741
Db	408	AGAAAGGAGGTATCCCTGTGACCAAGGCTTCAGGCTCTTCCCTGCACTCAGGACCA <td>349</td>	349
QY	1742	AGCCAGCTCACTCTGGGACCTGTGTTCCAGCATCTCTGTGCTCTGATTAAGATTC <td>1801</td>	1801
Db	348	AGCCAGCTCACTCTGGGACCTGTGTTCCAGCATCTCTGTGCTCTGATTAAGATTC <td>289</td>	289
QY	1802	TCCCTCCAGGCTTACGCTGGGATTTGGGACAGATTAAGATCCAACTATAGGCTAG <td>1861</td>	1861
Db	288	TCCCTCCAGGCTTACGCTGGGATTTGGGACAGATTAAGATCCAACTATAGGCTAG <td>229</td>	229
QY	1862	TTCCTGTCTAACTCAAGACTGTTCTGAGATGAGGCTCCAGGCTGTCAACCATG <td>1921</td>	1921
Db	228	TTCCTGTCTAACTCAAGACTGTTCTGAGATGAGGCTCCAGGCTGTCAACCATG <td>169</td>	169
QY	1922	CTGACCTGACCAAGGTTGAGGACATTAAGGAGGCTGTCTCTGTGGGACACTG <td>1981</td>	1981
Db	168	CTGACCTGACCAAGGTTGAGGACATTAAGGAGGCTGTCTCTGTGGGACACTG <td>109</td>	109
QY	1982	GAAAGTCCAGGTTGGGACCTCTGAGGAGCACTTGGGATCCCAATCCAGGTCATCT <td>2041</td>	2041
Db	108	GAAAGTCCAGGTTGGGACCTCTGAGGAGCACTTGGGATCCCAATCCAGGTCATCT <td>49</td>	49
QY	2042	CTAGGTTTGGATAC-CATAGTATGTATGTTTAACTGTGC 2081	
Db	48	CTAGGTTTGGATAC-CATAGTATGTATGTTTAACTGTGC 2081	
RESULT 8			
EX443303	948 bp	mRNA	linear
LOCUS	EX443303 Homo sapiens B CELLS (RAMOS CELL LINE)	Homo sapiens cDNA	EST 22-MAY-2003
DEFINITION	Clone CS0D007YF05 5'-PRIME, mRNA sequence.		
ACCESSION	EX443303		
VERSION	EX443303.1	GI:31018564	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
	1 (bases 1 to 948)		
	14, M.B., Gruber, C., Jesse, J. and Polyes, D.		
	Full-length cDNA libraries and normalization		
	Unpublished (2001)		
REFERENCE	Contact: Genoscope		
AUTHORS	Genoscope - Centre National de Sequencage		
TITLE	BP 191 91006 EVRY cedex - France		
JOURNAL	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr		
COMMENT	Library was constructed by Life Technologies, a division of		
	Invitrogen. This sequence belongs to sequence cluster 10000.f For		
	more information about this cluster, see		
	http://www.genoscope.cns.fr/		

```
cg1-bin/cluster.cgi?seg=CS0DGD007CC03QPL&cluster=10000.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DGD007CC03QPL.
Location/Qualifiers
1..948
source
sequence="cg1-bin-10000.gamians"
```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DG007YF05"
/tissue_type="B CELLS (RAMOS CELL LINE)"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-ctg150(47) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

```

ORIGIN

Query Match	41.5%;	Score 878.6;	DB 13;	Length 948;
Best Local Similarity	98.0%;	Pred. No. 5.1e-157;		
Matches 917;	Conservative	2;	Mismatches 14;	Indels 3;
				Gaps 3;

QY	21	CCGCTGCTAGAACCCCGGCGAAGGCGTGGAGCTGGGACTCCGAGCTCGGACAGACG	80
Db	15	CCGGTCCGAATTCCCCGGAGTGGCTGGAGCTGGGCTGGAGTCCGAGCTCGGACAGCG	74
QY	81	CA-CGGGCGGGCCCACTGGCTGGTGCCTCGAGGACTCTGAGCCCGCGGCGCCCGGCGC	139
Db	75	CAGNAGNCCGGCCCACTGGCTGGTGGCTCTGAGGCTCTGAACCCCGGCGGCGCCCGGCGC	134
QY	140	CACGCGGAACGACGGGAGCGAATGCGACCAACCCCTCTGGCTGCTCTCGCGGGTTCCCTG	199
Db	135	CACGCGGAACGACGGGAGCGAATGCGACCAACCCCTCTGGCTGCTCTCGCGGGTTCCCTG	194
QY	200	TCCAGAGAAGAACGGTGGATGGATTGGAATGACAACTTAGATACCGAGCGTCCCGTACAGAAA	259
Db	195	TCCAGAGAAGAACGGTGGATGGATTGGAATGACAACTTAGATACCGAGCGTCCCGTACAGAAA	254
QY	260	CGAGGCTCGAAGTGGGCGCCGAGCCAGACTGCGCCCTCGCTGTTGCCCTGAGCCCACT	319
Db	255	CGAGGCTCGAAGTGGGCGCCGAGCCAGACTGCGCCCTCGCTGTTGCCCTGAGCCCACT	314
QY	320	ACTGCTCCAGATCGTGCAACTGCTGTGGCACTGCTCCCGTCTTTGGGCCCTATATGCTTC	379
Db	315	ACTGCTCCAGATCGTGCAACTGCTGTGGCACTGCTCCCGTCTTTGGGCCCTATATGCTTC	374
QY	380	CTGGAGCCCGGAGAGGGCGGGCGGCGTAAACAGGGCCCTGSCATGAGCTACAGGCACTGAG	439
Db	375	CTGGAGCCCGGAGAGGGCGGGCGGCGTAAACAGGGCCCTGSCATGAGCTACAGGCACTGAG	434
QY	440	TATACCTGCAAGGTATACCCGCTCAAGAGAGCCTTGGCGGTGTGGAGCCCTTACCGCGCG	499
Db	435	TATACCTGCAH-GTGTACCCCGTCAAGAGAGCCTTGGCGGTGTGGAGCCCTTACCGCGCG	493
QY	500	CTGCCCCCGCAAGCATGTGCTCGGCCCACTGAGGTCCTGTGCTGTATCCAGGCTCCGC	559
Db	494	CTGCCCCCGCAAGCATGTGCTCGGCCCACTGAGGTCCTGTGCTGTATCCAGGCTCCGC	553
QY	560	TAGGCTTTTCACTCGAGCCCATGGGAGCATGSCACAGGCTGTGGAAAGCGCCACCGT	619
Db	554	TAGGCTTTTCACTCGAGCCCATGGGAGCATGSCACAGGCTGTGGAAAGCGCCACCGT	612
QY	620	ATCCCTGAGCTGAGGCTGCGGTCTTTCGCGCAGATGGCCACCGCCCTGGCGCACTGT	679
Db	613	ATCCCTGAGCTGAGGCTGCGGTCTTTCGCGCAGATGGCCACCGCCCTGGCGCACTGT	672
QY	680	CACGAGCACGCTGTGCTCGTGTGATCTCAAGCTGTGTGCTTTGTCTTGGCTGACCGT	739
Db	673	CACGAGCACGCTGTGCTCGTGTGATCTCAAGCTGTGTGCTTTGTCTTGGCTGACCGT	732
QY	740	GAGAGGAAGAGCTGTGCTGGAGAACCTGGAAGACTCTGCGGTGACTGGGCGAGAT	799

Db 733 GAGAGAAAGTGTGTCTGAGAACCTGAGAGACTCTCTGCTGTGACTGGGCCAGAT 792

QY 800 GATTCCTGTGGGACAGACAGCGGTGCGCAGCCTAAGTGGAGCCTGAGTACTCAGCTTA 859

Db 793 GATTCCTGTGGGACAGACAGCGGTGCGCAGCCTAAGTGGAGCCTGAGTACTCAGCTTA 852

QY 860 CGGGCTCTATCTCGGGGAGAGGAGCGATGTCTGAGCCTGGGCGTGGCGCTTTCACC 919

Db 853 CGGGCTCTATCTCGGGGAGAGGAGCGCGATGTCTGAGCCTGGGCGTGGCGCTTTCACC 912

QY 920 ATGCTGGCGGCGACCTACCCCTTCCAGAGCTGGAG 955

Db 913 ATGCTGGCGGCGACCTACCCCTTCCAGAGCTGGAG 948

RESULT 9
AL556690 1201 bp mRNA linear EST 31-MAY-2003
LOCUS CDNA clone CS0DK005YC05 5-PRIME, mRNA sequence.
ACCESSION AL556690.2 GI:31278491
VERSION EST.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1201)
AUTHORS Li, W.-B., Gruber, C., Jessee, J. and Polayres, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:12895955.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 10000.f For more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DK005AB03QPI&cluster=10000.f. Contact : Feng liang Email : fliang@life.techn.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DK005AB03QPI.
Location/Qualifiers
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK005YC05"
/cell_line="HELA"
/cell_line="HELA CELLS COT 25-NORMALIZED"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 41.2%; Score 872.4; DB 9; Length 1201;
Best Local Similarity 95.8%; Pred. No. 7,9e-156;
Matches 909; Conservative 15; Mismatches 22; Indels 3; Gaps 3;

QY 39 CAGGCTGAGCTGGCTTGGATCCGAGCTCGGAGCGAGCGAGCGGCCCACT 97

Db 63 CGGATTTGAGCTGGCTTGGATCCGAGCTCGGAGCGAGCGAGCGGCCCACT 122

QY 98 GGTGTGCTTGGAGGCTTGAAGCCCGGCGGCGGCGGCGGCGGAGAGAGAGGAGG 157

Db 123 GCTGTGCTTGGAGGCTTGAAGCCCGGCGGCGGCGGCGGCGGAGAGAGAGGAGG 182

QY 158 GAGATGCGAGCAACCCCTGTGCTCTCTGTGGGGTTCCTGTCAGAGAAAGCGTTG 217

Db 183 GAGATGCGAGCAACCCCTGTGCTCTCTGTGGGGTTCCTGTCAGAGAAAGCGTTG 242

QY 218 GAGTGTGATGACAACTTATGATATCCAGAGGCTCCCGTCAAGAAAGAGTGGAGGCC 277

Db 243 GAGTGTGATGACAACTTATGATATCCAGAGGCTCCCGTCAAGAAAGAGTGGAGGCC 302

QY 278 CAGCCCAAGTCCCGCTGCTGCTTGGCTTCCCTGAGCCCACTAAGTCTCAAGTGTGA 337

Db 303 CAGCCCAAGTCCCGCTGCTGCTTGGCTTCCCTGAGCCCACTAAGTCTCAAGTGTGA 362

QY 338 ACTGCTGTGGGCACTGCTCCCGTGTGGGCGCTATATGCTCTCTGAGCCCGAGAGGAG 397

Db 363 ACTGCTGTGGGCACTGCTCCCGTGTGGGCGCTATATGCTCTCTGAGCCCGAGAGGAG 422

QY 398 GGGCGGGCTTCAAGGCGCTTCACTGCTTCAAGGCACTGAGTATCTGCAAGGTATAC 457

Db 423 GGGCGGGCTTCAAGGCGCTTCACTGCTTCAAGGCACTGAGTATCTGCAAGGTATAC 482

QY 458 CCGCTTCAAGGAGGCTTGGCTGCTGCTGAGGCTTCAAGGCGGCTGCGCCCGACAGCAT 517

Db 483 CCGCTTCAAGGAGGCTTGGCTGCTGCTGAGGCTTCAAGGCGGCTGCGCCCGACAGCAT 542

QY 518 GTGCTGCGGCGCACTGAGGCTCTGCTGCTGATACCAAGCTCTCTAAGCTTTTCACTGG 577

Db 543 GTGCTGCGGCGCACTGAGGCTCTGCTGCTGATACCAAGCTCTCTAAGCTTTTCACTGG 602

QY 578 ACCCATGAGGAGCATGCAACAGCTGTGTGAGAGCGGCGGCGGCGGCGGAGGCT 637

Db 603 ACCCATGAGGAGCATGCAACAGCTGTGTGAGAGCGGCGGCGGCGGCGGAGGCT 661

QY 638 GCGGTGTCTCTCGGCGGAGAGGCGCACTGCTGCGGCGGCGGCGGCGGCGGCGGCT 697

Db 662 GCGGTGTCTCTCGGCGGAGAGGCGCACTGCTGCGGCGGCGGCGGCGGCGGCGGCT 721

QY 698 CTCGCTGATCTCAAGCTGTGTGCTTGTGCTTGTGCTGAGACCGTGAAGAGAGAGTGTG 757

Db 722 CTCGCTGATCTCAAGCTGTGTGCTTGTGCTTGTGCTGAGACCGTGAAGAGAGAGTGTG 781

QY 758 CTGAGAACTCTGAGAGACTCTCTGCTGCTGCTGAGGCGGCGGCGGCGGCGGCGGCT 817

Db 782 CTGAGAACTCTGAGAGACTCTCTGCTGCTGCTGAGGCGGCGGCGGCGGCGGCGGCT 841

QY 818 CAGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 877

Db 842 CAGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 901

QY 878 AAGGCAAGCGATGTCTGAGGCTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 937

Db 902 AAGGCAAGCGATGTCTGAGGCTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 961

QY 938 CCGTTCAGAGCTCGAGGCTGTCTGCTCTTTCGAGAGATCGGCGGCGGCGGCGGCT 986

Db 962 CCGTTCAGAGCTCGAGGCTGTCTGCTCTTTCGAGAGATCGGCGGCGGCGGCGGCGGCT 1009

RESULT 10
BG575275 975 bp mRNA linear EST 10-APR-2001
LOCUS 602597925F1 NIH_MGC_87 Homo sapiens CDNA clone IMAGE:4706575 5',
DEFINITION mRNA sequence.
ACCESSION BG575275
VERSION BG575275.1 GI:13582928
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 975)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.

Email: c9abbs@mail.nih.gov
 Tissue Procurement: DCTD/DRP
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed By: The I.M.A.G.E. Consortium (ULNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ULNL at:
 http://image.jnl.gov
 Place: LMAN0575 row: a column: 08
 High quality sequence stop: 891.
 Location/Qualifiers

FEATURES

source

1..975
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4706575"
 /issue_type="mammary adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 87"
 /note="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.383 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC library."

ORIGIN

Query Match 41.1%; Score 868.8; DB 12; Length 975;
 Best Local Similarity 96.2%; Pred. No. 3.8e-155;
 Matches 922; Conservative 0; Mismatches 32; Indels 4; Gaps 3;

QY 825 GCCCAGCTCTGCTGGGACCTGAGATCTCAGCTCAGCGGCTCATCTCGGAGCAAGCAG 884
 Db 1 GCCCAGCTCTGCTGGGACCTGAGATCTCAGCTCAGCGGCTCATCTCGGAGCAAGCAG 60
 QY 885 CCGATGCTCTGAGACCTGGGCGTGGCGCTTTCAACATCTGCGCGGCTCACTACCTCTCC 944
 Db 61 CCGATGCTCTGAGACCTGGGCGTGGCGCTTTCAACATCTGCGCGGCTCACTACCTCTCC 120
 QY 945 AGGACTGGAGCT 1004
 Db 121 AGGACTGGAGCT 180
 QY 1005 CAGGCTCTCTGCGGCTCTGCGGCTCTGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 1064
 Db 181 CAGGCTCTCTGCGGCTCTGCGGCTCTGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
 QY 1065 AACGCTCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1124
 Db 241 AACGCTCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 300
 QY 1125 TAGCCCCCAACCCGATCT 1184
 Db 301 TAGCCCCCAACCCGATCT 360
 QY 1185 TGGACGAGCCAGGAG 1244
 Db 361 TGGACGAGCCAGGAG 420
 QY 1245 CTACTACAGCGCTCAGCTCCCAACAGTGAATGAGTTTGGGGGTACTCTCAAGCCTTCTCC 1304
 Db 421 CTACTACAGCGCTCAGCTCCCAACAGTGAATGAGTTTGGGGGTACTCTCAAGCCTTCTCC 480
 QY 1305 TGCCCTCTGAACCTGAGCCAAACCTTCTGAGTCTCTCAGAGGAGAGAGAGAGAGAGAGAG 1364
 Db 481 TGCCCTCTGAACCTGAGCCAAACCTTCTGAGTCTCTCAGAGGAGAGAGAGAGAGAGAGAG 540
 QY 1365 GTGAGAGTGTCTGTATACATCTGCTTTGTTCACACACATGACAGTCTCTCTGGGTG 1424
 Db 541 GTGAGAGTGTCTGTATACATCTGCTTTGTTCACACACATGACAGTCTCTCTGGGTG 600
 QY 1425 CTATATCAGGTGCCAAGCCTCTGTTCTCGTGTCTGGAGTTCAGCAGTGAAGAGAGAGAG 1484
 Db 601 CTATATCAGGTGCCAAGCCTCTGTTCTCGTGTCTGGAGTTCAGCAGTGAAGAGAGAGAG 660

QY 1485 ATATTCCTCTCTCAGAGAGATGACAACTGGGATCTCTGAGCTGACACACTTTTCATG 1544
 Db 661 ATATTCCTCTCTCAGAGAGATGACAACTGGGATCTCTGAGCTGACACACTTTTCATG 720
 QY 1545 ACCATAGGTCACTGTCTCAGCTGAGTACATTT--GTACAGATGTGGCTTCACTGATG 1602
 Db 721 ACCATAGGTCACTGTCTCAGCTGAGTACATTTGTGTACCAAGTGTGGCTTCACTGATG 780
 QY 1603 CTGTGTCTCAGGCACTCTGTCTGCAAGCAATCTCTTTCACAAACCAAGCTCTCTTT 1662
 Db 781 CTGTGTCTCAGGCACTCTGTCTGCAAGCAATCTCTTTCACAAACCAAGCTCTCTTT 840
 QY 1663 -GTATCTGTACCTTTTC-AGAGAAAGAGAGTATCTCTGTGSCAAAGGCTCCAGGCTC 1720
 Db 841 GGTATCTGTACCTTTTCAGAGTAAGGAGAGAAATCTCTGGGCAAGGCTTCCAGGCTC 900
 QY 1721 TCCCTCGAAGCTCAGGACCCAGGACCCAGCTCTCAGTCTGAGATCTGTCTCCAGCATCT 1778
 Db 901 TCCCTCGAAGCTCAGGACCCAGGACCCAGCTCTGAGATCTGTCTCCAGCATCT 958

RESULT 11
 AL562576/c 1001 bp mRNA linear EST 31-MAY-2003
 LOCUS AL562576 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
 DEFINITION cDNA clone CS0DC013Y021 3-PRIME, mRNA sequence.

ACCESSION AL562576
 VERSION AL562576.2 GI:31286588
 KEYWORDS EST.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1001)
 AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT On Feb 15, 2001 this sequence version replaced gi:1291134.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 10000.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DC013CE11NP1&cluster=10000.f. Contact :
 Feng Liang Email : fliang@life.techn.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Paraday Avenue Genoscope sequence ID : CS0DC013CE11NP1.

FEATURES

source

1..1001
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DC013Y021"
 /issue_type="NEUROBLASTOMA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 40.8%; Score 862.4; DB 9; Length 1001;
 Best Local Similarity 94.0%; Pred. No. 6.3e-154;
 Matches 942; Conservative 33; Mismatches 18; Indels 9; Gaps 8;

QY 1081 AAGCATCTCTCTGACACCCCTGCTGCGACAGAGACCCCATGCTTACGCCCAACCGGATC 1140
 Db 1000 AAGCATCTCTCTGACACCCCTGCTGCGAGAGACCCCTATGCTTACGATCCAAACCGATC 941

```
QY 1141 CCATCTCTGGAGGCTGCCAGGTGTCCTCTGATGACTGGGGCTGGAGCAAGCCAGGGA 1200
Db 940 YVATCTTCTGGAGGCTGCCAGGTGTCCTCTGATGACTGGGGCTGGAGCAAGCCAGGGA 881
QY 1201 AAGAGGAGGAGACAGAGAGTGGTCTGTATGGCTAGACCAACCTTACTACACGCTCAGC 1260
Db 880 AAGAGGAGGAGACAGAGAGTGGTCTGTATGGCTAGACCAACCTTACTACACGCTCAGC 821
QY 1261 TGGCAACAGTGAATTGAGTTGGGGGTACTCCAGCCTTCTCCGCTCTGAACCTGAGC 1320
Db 820 TGCACACAGTGAATTGAGTTGGGGGTACTCCAGCCTTCTCCGCTCTGAACCTGAGC 762
QY 1321 GAAACCTTAGTGCCTTCCAGAGGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 761 GAAACCTTAGTGCCTTCCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 702
QY 1381 ACACATCTGCTTGTGTCCACACA-CATGAGTTCCTGCTGGGTGCTTATCGGTGCCAA 1439
Db 701 ACACATCTGCTTGTGTCCACACA-CATGAGTTCCTGCTGGGTGCTTATCGGTGCCAA 642
QY 1440 GCCCTGTCTGGTCTGGAGTACAGCACTGAGCAAGAGAGAGAGAGAGAGAGAGAG 1498
Db 641 GCCCTGTCTGGTCTGGAGTACAGCACTGAGCAAGAGAGAGAGAGAGAGAGAGAGAG 582
QY 1499 CAGAGATGACAAACATG-GCATCTCTGAGCTGACCAACTCTTTCATGACCAATAGGTC-AC 1556
Db 581 CAGAGATGACAAACATGCTGATCTCTGAGCTGACCAACTCTTTCATGACCAATAGGTC-AC 522
QY 1557 TGTCTACACTGGGTACACTTGTGACAGTGTCCGCTCCACGATGCTGCTGCTCAGGCA 1616
Db 521 TGTCTACACTGGGTACACTTGTGACAGTGTCCGCTCCACGATGCTGCTGCTCAGGCA 462
QY 1617 CCTCTGTCAGAGCAATCCCTTTCACAAACAAACAGCTGCC--TTTGATCTTGTACC 1674
Db 461 CCTCTGTCAGAGCAATCCCTTTCACAAACAAACAGCTGCC--TTTGATCTTGTACC 402
QY 1675 TTTTCAGAGAGAGAGAGATCCCTGTCGCCAAAGCTCTCAGGCTCTCCCTGCACTCA 1734
Db 401 TTTTCAGAGAGAGAGAGATCCCTGTCGCCAAAGCTCTCAGGCTCTCCCTGCACTCA 342
QY 1735 GAGACCAAGCCAGCTCACTCTGGGAATGTGTCCAGACACTCTCTGCTCTGATTA 1794
Db 341 GAGACCAAGCCAGCTCACTCTGGGAATGTGTCCAGACACTCTCTGCTCTGATTA 282
QY 1795 GAGATTTCTCTTCCAGGCTTAAGCTTGGGATTTGGGGCAGAGATAGAGATCCAACTATG 1854
Db 281 GAGATTTCTCTTCCAGGCTTAAGCTTGGGATTTGGGGCAGAGATAGAGATCCAACTATG 222
QY 1855 AGGCTAGTCTTGTCTAATCACTCAAGACTGTCTGGAATGAGGGT-CCAGGCTGTCAACA 1913
Db 221 AGGCTAGTCTTGTCTAATCACTCAAGACTGTCTGGAATGAGGGT-CCAGGCTGTCAACA 162
QY 1914 TGGGGCTTCTGACCTGAGCAACAAGGTGAGGAGAGAGATTAGGAGGGCTCTGCTGG 1973
Db 161 TGGGGCTTCTGACCTGAGCAACAAGGTGAGGAGAGAGATTAGGAGGGCTCTGCTGG 102
QY 1974 GCCACCTGAGAAATGCCAGTGGGACTCTTCTGGGAGCACTTGGGCTGCACAATCCAGG 2033
Db 101 GCCACCTGAGAAATGCCAGTGGGACTCTTCTGGGAGCACTTGGGCTGCSMAATCCGMA 42
QY 2034 TCCATCTAGTGTGATGATACCATGATGATGATGATGATGATGATGATGATGATGATG 2075
Db 41 CCNNNNNNNTAGGTTT-CAATCAGATGATGATGATGATGATGATGATGATGATGATGATG 1

RESULT 12
LOCUS BQ216198 873 bp mRNA linear EST 02-MAY-2002
DEFINITION AGENCOURT_7574831 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6059905
ACCESSION BQ216198
VERSION BQ216198.1 GI:20397598
KEYWORDS EST.
```

```
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-riemail.nih.gov
Tissue Procurement: DCTD/DTF/Gazdar
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed By: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM13327 row: f column: 02
High quality sequence scop: 640.
FEATURES
source
1..873
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6059905"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH MGC 68"
/note="Organ: lung; Vector: pCMV-Sport6; Site_1: NotI,
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.8 kb. Library constructed by Life
Technologies."
```

ORIGIN

Query Match 40.2%; Score 849.8; DB 13; Length 873;
Best Local Similarity 99.1%; Pred. No. 1.6e-151;
Matches 865; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

```
QY 111 AGGCTCTGAGCCCGCGGCGGCCCGGAGCCAGCAAGGAGGAGGAGGAGGAGGAGGAG 170
Db 1 AGGCTCTGAGCCCGCGGCGGCCCGGAGCCAGCAAGGAGGAGGAGGAGGAGGAGGAGGAG 60
QY 171 CCCCTGGGCTGCTCTGCGGGGTTCCTCTGTCAGAGAGAGAGCGGTTGAGTTGATGACA 230
Db 61 CCCCTGGGCTGCTCTGCGGGGTTCCTCTGTCAGAGAGAGAGCGGTTGAGTTGATGACA 120
QY 231 ACTTAGATACCGAGCGTCCCGTCCAGAAACAGCTGGAAGTGGGCCCCCAGCCAGATGCG 290
Db 121 ACTTAGATACCGAGCGTCCCGTCCAGAAACAGCTGGAAGTGGGCCCCCAGCCAGATGCG 180
QY 291 CCCCTGCTGTTGGCCCTGAGGCCCACTGCTCCAGATCGTGAACCTGTGGGCA 350
Db 181 CCCCTGCTGTTGGCCCTGAGGCCCACTGCTCCAGATCGTGAACCTGTGGGCA 240
QY 351 CTGCTCCCGGTCTGGGGCCCTATGCTCTCTGAGGCCAGAGAGGGCGGGCGGCTTACC 410
Db 241 CTGCTCCCGGTCTGGGGCCCTATGCTCTCTGAGGCCAGAGAGGGCGGGCGGCTTACC 300
QY 411 AGGCCCTGACAGCTCTACAGCAGTGAATATCTGCAAGGATGATCCCGTCCAGGAAG 470
Db 301 AGGCCCTGACAGCTCTACAGCAGTGAATATCTGCAAGGATGATCCCGTCCAGGAAG 360
QY 471 CCTTGGCGGTGTGAGAGCCCTTACGCGGGCTGCCCCCGCAGCAAGATGTGGCTCGGCCA 530
Db 361 CCTTGGCGGTGTGAGAGCCCTTACGCGGGCTGCCCCCGCAGCAAGATGTGGCTCGGCCA 420
QY 531 CTGAGGTCTGTGCTGTACCCAGCTCTCTAGGCTTTTTCCTGCTGAGACCTATGGGAGCA 590
Db 421 CTGAGGTCTGTGCTGTACCCAGCTCTCTAGGCTTTTTCCTGCTGAGACCTATGGGAGCA 480
QY 591 TGCAAGAGCTGTGTGGAAGCGGCCACGATATCCCTGAGCTTGAAGGCTCGGCTCTTCC 650
Db 481 TGCAAGAGCTGTGTGGAAGCGGCCACGATATCCCTGAGCTTGAAGGCTCGGCTCTTCC 540
```

QY 651 GCCAGATGCGCACCGCCCTGGCGCACTGTACACGACGAGTGTGCTGCTGATCTCA 710
 DB 541 GGCATATGGCCACCGCCCTGGCGCACTGTACACGACGAGTGTGCTGCTGATCTCA 600
 QY 711 AGCTGTGCTGCTTTGTCTTGTGCTGACCGTGAAGAGAAAGTGTGCTGAGAACTGG 770
 DB 601 AGCTGTGCTGCTTTGTCTTGTGCTGACCGTGAAGAGAAAGTGTGCTGAGAACTGG 660
 QY 771 AGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 830
 DB 661 AGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 QY 831 CCTAGCTGGAGCCTGAGATCACTCACTCACTCACTCACTCACTCACTCACTCACT 890
 DB 721 CCTAGCTGGAGCCTGAGATCACTCACTCACTCACTCACTCACTCACTCACTCACT 780
 QY 891 TCT-GAGCCTGGAGCCTGAGATCACTCACTCACTCACTCACTCACTCACTCACT 949
 DB 781 TCTGGAGCCTGGAGCCTGAGATCACTCACTCACTCACTCACTCACTCACTCACT 840
 QY 950 TGGAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 982
 DB 841 TGGAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 873

RESULT 13
 BQ941789 881 bp mRNA linear EST 21-AUG-2002
 LOCUS BQ941789
 DEFINITION AGENCOURT_8821009 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6422656
 5', mRNA sequence.

ACCESSION BQ941789
 VERSION BQ941789.1 GI:22357267
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE NIH-MGC http://mgi.nci.nih.gov/
 1 (bases 1 to 881)
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cga@rs-remail.nih.gov
 Tissue Procurement: DCTD/DTF/Gazdar
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LICM259 row: P column: 17
 High quality sequence stop: 689.
 Location/Qualifiers

FEATURES
 source
 1. 881
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6422656"
 /tissue_type="large cell carcinoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="NIH_MGC_18"
 /note="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming, directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GAGCGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."

ORIGIN
 Query Match 40.0%, Score 846.2, DB 13, Length 861;

Best Local Similarity 98.6%; Pred. No. 7.6e-151;
 Matches 863; Conservative 0; Mismatches 10; Indels 2; Gaps 1;

QY 643 GCTCTTCGCGCAGATGAGCGACCGCTTGGCGCATGTCAACGACGAGTGTGCTTGG 702
 DB 1 GCTCTTCGCGCAGATGAGCGACCGCTTGGCGCATGTCAACGACGAGTGTGCTTGG 60
 QY 703 TGATCTCAAGCTGTGCTGCTTGTCTTGTGCTGACCGTGAAGAGAAAGTGTGCTG 762
 DB 61 TGATCTCAAGCTGTGCTGCTTGTCTTGTGCTGACCGTGAAGAGAAAGTGTGCTG 120
 QY 763 GAACCTGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 822
 DB 121 GAACCTGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 QY 823 GTGCCAGCTCACTGAGAGACTGAGATCTGAGTCACTGAGTCACTGAGTCACTGAG 882
 DB 181 GTGCCAGCTCACTGAGAGACTGAGATCTGAGTCACTGAGTCACTGAGTCACTGAG 240
 QY 883 AGCCGATGTCTGAGAGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTT 942
 DB 241 AGCCGATGTCTGAGAGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTT 300
 QY 943 CGAGAGCTGAGAGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTTGG 1002
 DB 301 CGAGAGCTGAGAGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTTGG 360
 QY 1003 TGCAGGCTCTGAGGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTT 1062
 DB 361 TGCAGGCTCTGAGGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTT 420
 QY 1063 TGAACGCTCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1122
 DB 421 TGAACGCTCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 480
 QY 1123 CTAGGCTCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1182
 DB 481 CTAGGCTCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 540
 QY 1183 GCTGAGCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1242
 DB 541 GCTGAGCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 600
 QY 1243 CCTTATCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1302
 DB 601 CCTTATCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 660
 QY 1303 CCTGCTTGAAGCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1362
 DB 661 CCTGCTTGAAGCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 720
 QY 1363 GTGTGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1422
 DB 721 GTGTGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
 QY 1423 TGCTTATCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1482
 DB 781 TGCTTATCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 QY 1483 CAATATATCC--TGCTCAGAGATGACCAACTGG 1515
 DB 841 CAATATATCCCTGCTTCAAGATGACCAACTGG 875

RESULT 14
 BQ223497 918 bp mRNA linear EST 02-MAY-2002
 LOCUS BQ223497
 DEFINITION AGENCOURT_7576910 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6058986
 5', mRNA sequence.
 ACCESSION BQ223497
 VERSION BQ223497.1 GI:20404897
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 918)
 NIH-MGC <http://mgi.mgi.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs@mail.nih.gov
 Tissue Procurement: DCTD/DRP/Gazdar
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LIML at:
<http://image.liml.gov>
 Plate: LIML3324 row: 0 column: 19
 High quality sequence stop: 684.
 Location/Qualifiers

FEATURES

Source

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6058986"
 /tissue_type="large cell carcinoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 68"
 /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: oligo dT.
 Average insert size 1.8 kb. Library constructed by Life
 Technologies."

ORIGIN

Query Match 39.9%; Score 843.6; DB 13; Length 918;
 Best Local Similarity 98.5%; Pred. No. 2.4e-150;
 Matches 904; Conservative 0; Mismatches 9; Indels 5; Gaps 5;

63 CCGAGCTCGGACGACGCGCA-CGGGCGGGCCCACTGCTGTCCTGAGGGCTTACG 121
 1 CCGAGCTCGGACGACGCGCGGGCGGCCACCTGCTGTCCTGAGGGCTTACG 60

122 CCGGCGGGCGGGCGGGCGGCGGAAAGAGCGGGCGAGATCGAGCCACTTGGCT 181
 61 CCGGCGGGCGGGCGGGCGGCGGAAAGAGCGGGCGAGATCGAGCCACTTGGCT 120

182 GCTCCGCGGGGTTCCGTCAGGAAAGCGGTTGAGTGAATGACCACTTATGATCC 241
 121 GCTCTGCGGGGTTCCGTCAGGAAAGCGGTTGAGTGAATGACCACTTATGATCC 180

242 GAGCGTCCGTCGAGAAAGAGTCCGAGTGGGCGGCCGAGACTGCCCCCTGGCTG 301
 181 GAGCGTCCGTCGAGAAAGAGTCCGAGTGGGCGGCCGAGACTGCCCCCTGGCTG 240

302 TTGCCCCCTGAGCCACTTACTGCTCCAGATGTCGAACTGCTGTGGCACTGCTCCCT 361
 241 TTGCCCCCTGAGCCACTTACTGCTCCAGATGTCGAACTGCTGTGGCACTGCTCCCT 300

362 CTGGGCGCCCTATGCTCCCTGAGCGCGGAGGCGGGCGGGCTTACCAAGGCGCTGAC 421
 301 CTGGGCGCCCTATGCTCCCTGAGCGCGGAGGCGGGCGGGCTTACCAAGGCGCTGAC 360

422 TGGCCTTACAGGCACTAGATATACCTGCAAGGTGTACCCGTCAGAGAGCCCTGGCGTG 481
 361 TGGCCTTACAGGCACTAGATATACCTGCAAGGTGTACCCGTCAGAGAGCCCTGGCGTG 420

482 CTGAGACCTTACGCGGGCGGGCGGCGGCAAGACATGTGGCTGCGCCACTGAGTCTCG 541
 421 CTGAGACCTTACGCGGGCGGGCGGCGGCAAGACATGTGGCTGCGCCACTGAGTCTCG 480

542 GCTGTGATCCAGCTCTCTACAGCTTTTTCACCTGGAGCCATGAGGAGCAATGCAAGCTCG 601
 481 GCTGTGATCCAGCTCTCTACAGCTTTTTCACCTGGAGCCATGAGGAGCAATGCAAGCTCG 540

QY 602 GTGGAGACGCGCACCGTATCTCTGAGGCTTGAAGGCTGCGCTTTCGCCAGATGCGC 661
 DB 541 GTGGAGACGCGCACCGTATCTCTGAGGCTTGAAGGCTGCGCTTTCGCCAGATGCGC 600

QY 662 ACCGCGCTGGGCGCATGTCACAGACGAGTGTGCTGCTGCTGATCTCAACTGTGTGCG 721
 DB 601 ACCGCGCTGGGCGCATGTCACAGACGAGTGTGCTGCTGCTGATCTCAACTGTGTGCG 660

QY 722 TTGTCTTTCGCTGACCGGTGAGAGAGAGAGTGTGTGTGAGAACTGAGAGACTCTGCG 781
 DB 661 TTGTCTTTCGCTGACCGGTGAGAGAGAGAGTGTGTGTGAGAACTGAGAGACTCTGCG 720

QY 782 GTGTGACTGGGCGGAGATGATTTCTTGTGGGCAAGACAGCGTGGCCAGCTTACGTGGA 841
 DB 721 GTGTGACTGGGCGGAGATGATTTCTTGTGGGCAAGACAGCGTGGCCAGCTTACGTGGA 780

QY 842 CCTAGATATCTACAGTACGCGGCGCTCATCTCGGCGCAAGAGCGGAGTGTGTGAGGCTG 901
 DB 781 CCTAGATATCTACAGTACGCGGCGCTCATCTCGGCGCAAGAGCGGAGTGTGTGAGGCTG 840

QY 902 GCGG-TGGCGCTCTTTCACATGCT-GGCGGCGCACTACCCC-TTCCAGAGCTC-GGAGCC 957
 DB 841 GCGGTTGGCGCTCTTTCACATGCTGGGCGGCGCACTACCCCTTCCAGAGCTGGGAGCC 900

QY 958 TGTCTGCTCTTGTGGCA 975
 DB 901 TGTCTGCTCTTGTGGCA 918

RESULT 15

LOCUS

EX325345
 EX325345

DEFINITION

EX325345 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
 Homo sapiens cDNA clone CS0DL004YF04 5-PRIME, mRNA sequence.

ACCESSION

EX325345
 EX325345.1 GI:30338449

VERSION

EST.
 EST.

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 945)

AUTHORS

L.W.B., Gruber,C., Jesse,J. and Polyes,D.

TITLE

Full-length cDNA libraries and normalization

JOURNAL

Unpublished (2001)

COMMENT

Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: sequef@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 10000.f For
 more information about this cluster, see
<http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DL004D0020P1&cluster=10000.f>
 Feng Liang Email: fliang@life.com URL: <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID: CS0AL004D020P1.

FEATURES

Location/Qualifiers

1..945

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DL004YF04"
 /cell_type="B CELLS (RAMOS CELL LINE)"
 /cell_line="RAMOS CELL LINE"
 /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSport 6 vector. Library was normalized."

ORIGIN

PE 14-SEP-2000; 2000MO-US025515.
XX
PR 15-SEP-1999; 99US-0154141P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Yue H, Tang YT, Bandman O, Hillman JL, Baughn MK, Azimzai Y;
PL Lu DM;
XX
DR WPI; 2001-244811/25.
DR N-PSDB; AAF30480.
XX
XX
PT Novel human proein phosphatase and kinase proteins for diagnosis,
PT treatment and prevention of gastrointestinal, immune system, neurological
PT and cell proliferative disorders.
XX
XX
XX Claim 1; Page 88-89; 103pp; English.

CC The present sequence is that of novel human protein phosphatase and
CC kinase protein PPHKP-5, as predicted from Incyte Clone ID No. 1211505CB1
CC (see AAF030480). Tissues that express PPHKP-5 (as a fraction of total
CC tissue expressing PPHKP-5) include reproductive (0.288),
CC gastrointestinal (0.212) and haematopoietic or immune (0.192). Diseases
CC or conditions associated with tissues expressing PPHKP-5 (as a fraction
CC of total tissues expressing PPHKP-5) include cancer (0.577), inflammation
CC or trauma (0.327) and cell proliferation (0.308). The encoded protein
CC shows homology to rat kinase. The invention provides human PPHKP-1 to -11
CC polypeptides (see AAB20322-32) and polynucleotides (see AAF030476-86). It
CC also provides expression vectors, host cells, antibodies, agonists and
CC antagonists, as well as methods for diagnosing, treating or preventing
CC disorders associated with expression of PPHKP, including gastrointestinal
CC disorders, immune system disorders, neurological disorders and cell
CC proliferative disorders, including cancer

Query Match	100.0%	Score 1892;	DB 4;	Length 358;
Best Local	100.0%	Pred. No. 6.1e-176;		
Matches 358;	0;	Mismatches	0;	Gaps 0;
Conservative				

[illegible]

RESULT 2	
AAU03509	
XX	AAU03509 standard; protein; 358 AA.
XX	
AC	AAU03509;
XX	
DT	12-SEP-2001 (first entry)
XX	

DE	Human protein kinase #9.
XX	
KW	Human; protein kinase; PK; STK; cancer; cardiovascular disease;
KW	metabolic disorder; immune related disease; neurological disorder;
KW	neurodegenerative disorder; inflammatory disorder; infectious disease;
KW	reproductive disorder.

OS	Homo sapiens.
XX	
PN	WO200138503-A2.
XX	
PD	31-MAY-2001.

PF 22-NOV-2000; 2000WO-US032085.

PR 24-NOV-1999; 99US-0167482P.

PA (SUGE-) SUGEN INC.

PI Plowman GD, Whyte D,

XX
XX

DR N-PSDB; AAS06709.

PT Nucleic acids encoding human kinase polypeptides, useful for preventing

PT neuronal-associated diseases, and microbial infections.

PS Claim 7; Fig 2; 433pp; English.
xx

CC AU03501-AU03557 represent novel human protein kinases #1-57. The novel

serine/threonine kinase (PTK and STK) families. The polynucleotides

CC prevention, diagnosis and treatment of diseases associated with

CC cancers (especially cancers of haematopoietic origin), cardiovascular

CC immune related diseases (e.g. rheumatoid arthritis), neurological disorders (e.g. schizophrenia);

CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious

CC Additionally, polynucleotides encoding protein kinases may be used for

CC polypeptides may be used as antigens in the production of antibodies against the protein kinases and is essential for the

CC protein kinase expression and activity

Sequence 358 AA;

Query Match	99.8%	Score 1888;	DB 4;	Length 358;
Best Local Similarity	99.7%	Pred. No. 1.5e-175;		
Matches 357; Conservative	1;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	MAATPLAAAGSISKKKLELDNDNIDTERPVOKAASASPORLPCCILSPSPPAAPRAT	60
Db	1	MAATPLAAAGSISKKKLELDNDNIDTERPVOKAASASPORLPCCILSPSPPAAPRAT	60
QY	61	AAVATASRLGPVYLLEPEEGGAYOALHQPSTEXTKCYVPQGEALVLEPAARLPKHKV	120
Db	61	AAVATASRLGPVYLLEPEEGGAYOALHQPSTEXTKCYVPQGEALVLEPAARLPKHKV	120

Db 121 APTFVLTGTCVLAFFPRTRGDHMSLVRSRRRLPEPEAAVLFFQMTALAHCHQHELVL 1801

Qy 181 RULKLCRPVPAADREKKVLVLENLBDSCVLTGPDSDLMDKTACPAYVPEILSSPASYSGK 240

Db 181 RULKLCRPVPAADREKKVLVLENLBDSCVLTGPDSDLMDKTACPAYVPEILSSPASYSGK 240

Qy 241 PADVSLGVALLFTMLAGHYPRQDSPEVLTFKTRRGAYVALPAGLSAPARCLVCLLRREP 3001

DB 241 AADVMSLGVALFTMLAGHYPFODSEPVLLFGKIRRGAYALPAGLSAPARCLVRCILRRBP 300
 QY 301 AERLTATGILLHPMLRQDPMPLAPTRSHLMEAAQVVPDGLGDEAREEGDREVLVYG 358
 DB 301 AERLTATGILLHPMLRQDPMPLAPTRSHLMEAAQVVPDGLGDEAREEGDREVLVYG 358

RESULT 3
 AAM38908
 ID AAM38908 standard; protein; 358 AA.
 AC AAM38908;
 XX
 XX 22-OCT-2001 (first entry)
 DE Human polypeptide SEQ ID NO 2053.
 XX
 XX Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KM peripheral nervous system; neuropathy; central nervous system; CNS;
 KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemoclastic;
 KM chemokine; thrombolytic; drug screening; arthritis; inflammation;
 XX leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US034263.
 XX
 PR 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX
 XX (HYSE-) HYSBQ INC.
 PA
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX
 XX WPI; 2001-442253/47.
 DR N-PDB; AAI58064.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 XX
 PS Example 3; SEQ ID NO 2053; 10078bp; English.
 XX
 XX The invention relates to human nucleic acids (AA157798-AA161369) and the
 CC encoded polypeptides (AAM38642-AAM42213) with noctropic.
 CC immunosuppressant and cyostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression.
 CC Actin/limbin activity, chemoclastic/chemokine activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemia and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX
 SQ Sequence 358 AA;

Query Match 99.6%; Score 1885; DB 4; Length 358;
 Best Local Similarity 99.4%; Pred. No. 2.9e-175;
 Matches 356; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEATPLAPAGSLSRKRLLELDNDPTERRPYOKRARSQPOPLPCLPLSPPTAPDRT 60
 DB 1 MRATPLAPAGSLSRKRLLELDNDPTERRPYOKRARSQPOPLPCLPLSPPTAPDRT 60
 QY 61 AVATASRLGPLYLLEPEBGRAYQALHCPTEYTCYYPVQBALAVLEPYARLPKHV 120
 DB 61 AVATASRLGPLYLLEPEBGRAYQALHCPTEYTCYYPVQBALAVLEPYARLPKHV 120
 QY 121 ARPTEVLAGTQLLYAFETRTGDHMSLVTRTHRIPEPAAVLFRQMATLALHCHQGLVL 180
 DB 121 ARPTEVLAGTQLLYAFETRTGDHMSLVTRTHRIPEPAAVLFRQMATLALHCHQGLVL 180
 QY 181 RDLKICREVPADREKKLVLENTLSDSCVLTGDDSLMFKHCPAVGELLSSRASYSCK 240
 DB 181 RDLKICREVPADREKKLVLENTLSDSCVLTGDDSLMFKHCPAVGELLSSRASYSCK 240
 QY 241 AADVMSLGVALFTMLAGHYPFODSEPVLLFGKIRRGAYALPAGLSAPARCLVRCILRRBP 300
 DB 241 AADVMSLGVALFTMLAGHYPFODSEPVLLFGKIRRGAYALPAGLSAPARCLVRCILRRBP 300
 QY 301 AERLTATGILLHPMLRQDPMPLAPTRSHLMEAAQVVPDGLGDEAREEGDREVLVYG 358
 DB 301 AERLTATGILLHPMLRQDPMPLAPTRSHLMEAAQVVPDGLGDEAREEGDREVLVYG 358

RESULT 4
 AAM40694
 ID AAM40694 standard; protein; 393 AA.
 XX
 AC AAM40694;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 XX Human polypeptide SEQ ID NO 5625.
 DE
 XX
 XX Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KM peripheral nervous system; neuropathy; central nervous system; CNS;
 KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemoclastic;
 KM chemokine; thrombolytic; drug screening; arthritis; inflammation;
 XX leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US034263.
 XX
 PR 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX
 XX (HYSE-) HYSBQ INC.
 PA
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX
 XX WPI; 2001-442253/47.
 DR N-PDB; AAI59850.

XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
XX
PS Example 2; SEQ ID NO 5625; 10078bp; English.
XX
CC The invention relates to human nucleic acids (AA157798-AA161369) and the
CC encoded polypeptides (AA158642-AA162213) with neurotrophic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemocytic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 393 AA;

Query Match 99.4%; Score 1885; DB 4; Length 393;
Best Local Similarity 99.4%; Pred. No. 3.3e-175;
Matches 356; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRATPLAAPAGSLSRKKRLDNDLTERPVQKRAAGGQPPRLPCLPLSPPTADPRAT 60
DB 36 MRATPLAAPAGSLSRKKRLDNDLTERPVQKRAAGGQPPRLPCLPLSPPTADPRAT 95
QY 61 AVATASRLGPGYVLLPEEGGGRAYQALHCPGTGYTCVKVPVOEALAVLEPYARLPKHKAV 120
DB 96 AVATASRLGPGYVLLPEEGGGRAYQALHCPGTGYTCVKVPVOEALAVLEPYARLPKHKAV 155
QY 121 ARPEVLAGTQLLYAFPTRTGDMHSLVTRRRIPEPEAAVLFROMATLACHQHGVLV 180
DB 156 ARPEVLAGTQLLYAFPTRTGDMHSLVTRRRIPEPEAAVLFROMATLACHQHGVLV 215
QY 181 RDLKICRFEPADREKKVLVLENEEDSCVLTGPDSDLMDBACPAVYGPPISSRASYSYGK 240
DB 216 RDLKICRFEPADREKKVLVLENEEDSCVLTGPDSDLMDBACPAVYGPPISSRASYSYGK 275
QY 241 AADVMSLGVALLFTMLAGHYFPQDSSEVLLFGKIRRGAVYALPAGLSAPACLVRCCLRRBP 300
DB 276 AADVMSLGVALLFTMLAGHYFPQDSSEVLLFGKIRRGAVYALPAGLSAPACLVRCCLRRBP 335
QY 301 AERLTATGILLHPWLKQDPMPLAPTRSHLWEAAQVVPDGLGDEAREEGDEEVVLYG 358
DB 336 AERLTATGILLHPWLKQDPMPLAPTRSHLWEAAQVVPDGLGDEAREEGDEEVVLYG 393

RESULT 5
ADCS9337
ID ADCS9337 standard; protein; 358 AA.
AC ADCS9337;
XX
XX 18-DEC-2003 (first entry)
XX
XX Human cell death inhibitory protein.
XX
XX Human; cell-death; cardiac; neuroprotective; anti-HIV; antiinflammatory;
XX cerebroprotective; hepatotropic; cytostatic; immunosuppressive;
XX antirheumatic; ophthalmological; neurotropic; antiparkinsonian;
XX anticonvulsant; hypotensive; antiarteriosclerotic; haemostatic;
XX antialcoholic; virucide; HIV; cardiac disease; immunological disease;
XX neurodegenerative disease; ischaemic damage;
XX central nervous system disorder; hepatic disorder;
XX acute pancreatic inflammation; cancer; AIDS; autoimmune disease;
XX rheumatism; Crohn's disease; glaucoma; Alzheimer's disease;
XX Parkinson's disease; Huntington's disease; hypertension;

KW arteriosclerosis; reperfusion damage; myocardial infarction;
KW cerebral trauma; cerebral infarction; cerebral haemorrhage; hepatitis;
KW alcoholic hepatitis; cerebral ischemia.
XX
XX Homo sapiens.
XX
XX JP2003063986-A.
XX
XX
XX 05-MAR-2003.
XX
XX 27-AUG-2001; 2001JP-00255811.
XX
XX 27-AUG-2001; 2001JP-00255811.
XX
XX (SUMTOMO SEIYAKU KK.
XX WPI; 2003-601360/57.
XX
XX N-PSDB; ADCS9338.
XX
XX Controlling cell-death by administering positive or negative regulator of
PT cell-death inhibitory-factor, for treating cancer, AIDS, autoimmune
PT diseases, Crohn's diseases, glaucoma, Alzheimer disease.
XX
XX
XX Claim 1; SEQ ID NO 3; 35pp; Japanese.
XX
CC This invention relates to a novel method for controlling cell-death
CC comprising administering a positive or negative regulator of cell-death
CC inhibitory-factor. The invention also comprises a method for screening
CC modulators of cell-death, by contacting cells expressing cell-death
CC inhibitory factor with a candidate compound, monitoring level of
CC expression of cell-death inhibitory factor, evaluating cell-death
CC modulation ability of the compound based on change in the level of
CC expression of the factor and selecting compounds having cell-death
CC modulation ability. The cell death regulator of the invention may have
CC cardiac, neuroprotective, anti-HIV, antiinflammatory, cerebroprotective,
CC hepatotropic, cytostatic, immunosuppressive, antipneumatic,
CC ophthalmological, neurotropic, antiparkinsonia, anticonvulsant,
CC hypotensive, antiarteriosclerotic haemostatic, antialcoholic and
CC virucide activities. The method of the invention is useful for treating
CC HIV, cardiac diseases, immunological diseases, neurodegenerative disease,
CC ischaemic damage and congestion, disorder of central nervous system,
CC hepatic disorder, acute pancreatic inflammation, and cancer, AIDS,
CC autoimmune diseases, rheumatism, Crohn's disease, glaucoma, Alzheimer's
CC disease, Parkinson's disease, Huntington's disease, hypertension,
CC arteriosclerosis, reperfusion damage, myocardial infarction, cerebral
CC trauma, cerebral infarction, cerebral haemorrhage, hepatitis, alcoholic
CC hepatitis, and cerebral ischemia. The present sequence represents the
CC human cell death inhibitory protein of the invention.
XX
SQ Sequence 358 AA;

Query Match 99.4%; Score 1881; DB 7; Length 358;
Best Local Similarity 99.4%; Pred. No. 7.2e-175;
Matches 356; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRATPLAAPAGSLSRKKRLDNDLTERPVQKRAAGGQPPRLPCLPLSPPTADPRAT 60
DB 1 MRATPLAAPAGSLSRKKRLDNDLTERPVQKRAAGGQPPRLPCLPLSPPTADPRAT 60
QY 61 AVATASRLGPGYVLLPEEGGGRAYQALHCPGTGYTCVKVPVOEALAVLEPYARLPKHKAV 120
DB 61 AVATASRLGPGYVLLPEEGGGRAYQALHCPGTGYTCVKVPVOEALAVLEPYARLPKHKAV 120
QY 121 ARPEVLAGTQLLYAFPTRTGDMHSLVTRRRIPEPEAAVLFROMATLACHQHGVLV 180
DB 121 ARPEVLAGTQLLYAFPTRTGDMHSLVTRRRIPEPEAAVLFROMATLACHQHGVLV 180
QY 181 RDLKICRFEPADREKKVLVLENEEDSCVLTGPDSDLMDBACPAVYGPPISSRASYSYGK 240
DB 181 RDLKICRFEPADREKKVLVLENEEDSCVLTGPDSDLMDBACPAVYGPPISSRASYSYGK 240
QY 241 AADVMSLGVALLFTMLAGHYFPQDSSEVLLFGKIRRGAVYALPAGLSAPACLVRCCLRRBP 300
DB 241 AADVMSLGVALLFTMLAGHYFPQDSSEVLLFGKIRRGAVYALPAGLSAPACLVRCCLRRBP 300

DB 241 AADVWSLGVALLFTMLAGHYPPFODSEPVLLFGKIRGAYALPAGLSAPARCLVRCILRRPP 300
QY 301 AERLTATGILLHPMLRQDPMPLAPTRSHLWEAAQVYPDGLGDEAREEGDREVLYG 358
DB 301 AERLTATGILLHPMLRQDPMPLAPTRSHLWEAAQVYPDGLGDEAREEGDREVLYG 358
RESULT 6
ID ABB80976 standard; protein; 360 AA.
AC ABB80976;
XX
DT 21-OCT-2002 (first entry)
DE Human tribbles homologue-3 (htrb-3) polypeptide encoding DNA.
XX
KW Human; tribbles; htrb-1; stress kinase inhibitor protein; SKIP-1; AP-1;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic; osteopathic;
KW ophthalmological; cardiac; cytosolic; haemostatic; immunosuppressive;
KW antiinflammatory; estrogen receptor; fibroblast growth factor; FGF;
KW tumor necrosis factor; TNF; htrb-3.
XX
OS Homo sapiens.
PN WO200253743-A2.
XX
PD 11-JUL-2002.
XX
PF 08-JAN-2002; 2002WO-US000070.
XX
PR 08-JAN-2001; 2001US-0260294P.
XX
PA (INTE-) INTERLEUKIN GENETICS INC.
XX
PI Dower S, Qunastrom E, Kiss-Toch E;
XX WPI; 2002-590635/63.
DR N-PSDB; ABB86479.
XX
PT Novel isolated human tribbles homologue-1 polypeptide for inhibiting AP-1
PT -mediated inflammatory signal in a cell, and activating ERK-mediated
PT signal e.g. AP-1-mediated gene activation signal in a cell.
XX
PS Example 10; Fig 11B; 131pp; English.
XX
CC The invention provides an isolated human tribbles homologue-1 (htrb-1,
CC also known as stress kinase inhibitor protein (SKIP-1)) polypeptide. The
CC htrb polypeptide is useful for inhibiting an AP-1 mediated inflammatory
CC signal in a cell. The polypeptide employed in the method is preferably
CC htrb-1, htrb-1 N htrb-1 C, htrb-1 N C, htrb-3, htrb-3 N htrb-3 C, or htrb
CC -3 N C. It is also useful for providing htrb agonist activity for
CC activating an ERK-mediated signal e.g. AP-1-mediated gene activation
CC signal, an estrogen receptor-mediated gene activation signal, an
CC fibroblast growth factor (FGF) induced signal, or a PMA induced signal,
CC in a cell. Htrb modulators are useful for modulating AP-1 mediated
CC inflammatory signal in a cell such as tumor necrosis factor (TNF) induced
CC inflammatory signal, or an interleukin induced inflammatory signal. htrb
CC proteins are useful in screening assays, predictive medicine and in
CC compounds e.g. for treating and/or preventing diseases caused by abnormal
CC htrb activity, such as rheumatoid arthritis, diabetes, psoriasis,
CC osteoporosis, diabetic retinopathy, myocardial infarction and cancers.
CC The htrb therapeutics are useful for antagonizing interleukin-1 dependent
CC disorders of human placenta, intraventricular hemorrhage, neonatal white
CC matter damage and subsequent cerebral palsy, and inflammation or
CC autoimmune disorders. The present sequence represents the htrb-3
CC polypeptide
XX
SQ Sequence 360 AA;
Query Match 98.9%; Score 1871; DB 5; Length 360;
Best Local Similarity 98.6%; Pred. No. 6.9e-1/4;

Matches 355; Conservative 3; Mismatches 0; Indels 2; Gaps 1;
QY 1 MRATPLAAPAGSLSRKKRLIEDNDLTERPVQKARSGGQRRPLPLSPPTAPDQAT 60
DB 1 MRATPLAAPAGSLSRKKRLIEDNDLTERPVQKARSGGQRRPLPLSPPTAPDQAT 60
QY 61 AVATASRLGPPVLLPEPEGGAAYQALHCPCTCTCKYYPQVQALAVLEPARLPBPHGV 120
DB 61 AVATASRLGPPVLLPEPEGGAAYQALHCPCTCTCKYYPQVQALAVLEPARLPBPHGV 120
QY 121 ARPTVLGCTQLLYAFPTRTGDMSLVTRHRIPEPEAAVLFROMATALAHCHQGLVL 180
DB 121 ARPTVLGCTQLLYAFPTRTGDMSLVTRHRIPEPEAAVLFROMATALAHCHQGLVL 180
QY 181 RDLKLCRFVFA--DREKKVLLENLEDSCVLTGPDDSLMDKACPAVYGPETLSSRASYS 238
DB 181 RDLKLCRFVFA--DREKKVLLENLEDSCVLTGPDDSLMDKACPAVYGPETLSSRASYS 240
QY 239 GKADVWSLGVALLFTMLAGHYPPFODSEPVLLFGKIRGAYALPAGLSAPARCLVRCILRR 298
DB 241 GKADVWSLGVALLFTMLAGHYPPFODSEPVLLFGKIRGAYALPAGLSAPARCLVRCILRR 300
QY 299 BPAERLTATGILLHPMLRQDPMPLAPTRSHLWEAAQVYPDGLGDEAREEGDREVLYG 358
DB 301 BPAERLTATGILLHPMLRQDPMPLAPTRSHLWEAAQVYPDGLGDEAREEGDREVLYG 360
RESULT 7
ID AAY69157 standard; peptide; 360 AA.
XX
AC AAY69157;
XX
DT 30-MAY-2000 (first entry)
XX
DE Peptidea UJ503-KS comprising domains VIA to XI of a protein kinase.
XX
KW Kinase activity; molecular weight marker; isoelectric focusing marker;
KW peptide fragmentation control; cellular signal transduction.
XX
OS Homo sapiens.
PN WO2000008180-A2.
XX
PD 17-FEB-2000.
XX
PF 03-AUG-1999; 99WO-US017630.
XX
PR 04-AUG-1998; 98US-0095270P.
XX 11-SEP-1998; 98US-0099972P.
XX
PA (IMMV) IMMUNEX CORP.
XX
PI Virca GD, Bird TM, Anderson DM, Marken JS;
XX WPI; 2000-195584/17.
DR N-PSDB; AA61155.
XX
PT New human kinase polypeptides and polynucleotides used as molecular
PT weight markers and as controls for peptide fragmentation.
XX
PS Claim 2; Page 10; 60pp; English.
XX
CC The present sequence represents a partial polypeptide which has kinase
CC activity. The kinase polynucleotides can be used to express the
CC polypeptides, and as probes to identify nucleic acids encoding proteins
CC having kinase activity. The kinase polypeptides and isoelectric focusing
CC markers, and as controls for peptide fragmentation. They also have a
CC number of therapeutic uses as kinases play a central role in cellular
CC signal transduction. The polypeptides could also be used to identify
CC binding partner proteins. The polypeptides can also be used as a reagent
CC to identify any proteins that the polypeptide regulates, and proteins

CC with which it might interact. The polypeptides may also be used for
 CC preparation of antibodies. The antibodies can be used in assays to detect
 CC the presence of the protein, and to purify the protein by immunofluorescence
 CC chromatography

XX Sequence 360 AA:

Query Match 88.6%; Score 1676; DB 3; Length 360;
 Best Local Similarity 99.7%; Pred. No. 7,8e-155;
 Matches 317; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRATPLAAPAGSLSRKKRLLELDNLDTERPVQKRAAGPQRLPCLLPSPPTAPDRAT 60
 DB 43 MRATPLAAPAGSLSRKKRLLELDNLDTERPVQKRAAGPQRLPCLLPSPPTAPDRAT 102
 QY 61 AVATASRLGPGYVLLPEBEGRAYQALHCPGTGTYCKVYVQBALVLEBYARLPKHV 120
 DB 103 AVATASRLGPGYVLLPEBEGRAYQALHCPGTGTYCKVYVQBALVLEBYARLPKHV 162
 QY 121 ARPTVLATGTLVAFPTTHGDMHSLVTRHRIPEPEAAVLFROMATLAHQHGLV 180
 DB 163 ARPTVLATGTLVAFPTTHGDMHSLVTRHRIPEPEAAVLFROMATLAHQHGLV 222
 QY 181 RDLKLCRFVADREKKLVLENLEDSCVLTGPDSDMDKACPAVYGPETLSSRASYSK 240
 DB 223 RDLKLCRFVADREKKLVLENLEDSCVLTGPDSDMDKACPAVYGPETLSSRASYSK 282
 QY 241 AADWMSLGVALLFTMLAGHPFODSEVLLFGKIRRGAYALPAGLSAPARCLVCLRRBP 300
 DB 283 AADWMSLGVALLFTMLAGHPFODSEVLLFGKIRRGAYALPAGLSAPARCLVCLRRBP 342
 QY 301 AERLTATGILLHPWLROD 318
 DB 343 AERLTATGILLHPWLROD 360

RESULT 8
 AAB85791
 ID AAB85791 standard; protein; 323 AA.

AC AAB85791;
 DT 29-OCT-2001 (first entry)
 DE Human kinase PKIN-10.
 KM PKIN; kinase; cytosolic; immunosuppressive; immunostimulant; human;
 KM antiarteriosclerotic; cardiant; gene therapy; antisense therapy.
 OS Homo sapiens.
 OS WO200160991-A2.
 PN 23-ANG-2001.
 PD 16-FEB-2001; 2001WO-US005240.
 PF 17-FEB-2000; 2000US-0183682P.
 PR 02-MAR-2000; 2000US-0186559P.
 PR 09-MAR-2000; 2000US-0188606P.
 PR 17-MAR-2000; 2000US-0189989P.
 PR 30-MAR-2000; 2000US-0193851P.
 PA (INCY-) INCYTE GENOMICS INC.

PI Tang YT, Buford N, Gandhi AR, Patterson C, Khan JA, Yue H;
 PI Hafalia A, Shih LT, Tribouley CM, Yao MG, Burdill JD, Marcus GA;
 PI Zingler KA, Lu DAM, Bandman O, Policky JL, Griffin JA, Thornton M;
 PI Nguyen DB, Lai P, Walsh RT;
 DR WPI: 2001-514771/56.
 DR N-PSDB; AAH76218.
 XX

PT isolated human kinase polypeptides useful in the diagnosis, treatment and
 PT prevention of cancer, immune disorders and disorders affecting growth and
 PT development.

XX Claim 1; Page 115; 126pp; English.

CC The invention provides human kinases (PKIN) and polynucleotides encoding
 CC PKIN. The PKIN polypeptides can be expressed using standard recombinant
 CC methodology. The PKIN polypeptides, polynucleotides, modulators and
 CC specific antibodies are useful in the diagnosis, treatment and prevention
 CC of cancer, immune disorders, disorders affecting growth and development,
 CC atherosclerosis, and other cardiovascular diseases, and lipid disorders
 CC and in the assessment of the effects of exogenous compounds on the
 CC expression of nucleic acid sequences of human kinases. The present
 CC sequence represents a human PKIN-10 polypeptide

XX Sequence 323 AA:

Query Match 83.8%; Score 1585.5; DB 4; Length 323;
 Best Local Similarity 86.3%; Pred. No. 4.7e-146;
 Matches 309; Conservative 3; Mismatches 11; Indels 35; Gaps 2;

QY 1 MRATPLAAPAGSLSRKKRLLELDNLDTERPVQKRAAGPQRLPCLLPSPPTAPDRAT 60
 DB 1 MRATPLAAPAGSLSRKKRLLELDNLDTERPVQKRAAGPQRLPCLLPSPPTAPDRAT 60
 QY 61 AVATASRLGPGYVLLPEBEGRAYQALHCPGTGTYCKVYVQBALVLEBYARLPKHV 120
 DB 61 AVATASRLGPGYVLLPEBEGRAYQALHCPGTGTYCKVYVQBALVLEBYARLPKHV 120
 QY 121 ARPTVLATGTLVAFPTTHGDMHSLVTRHRIPEPEAAVLFROMATLAHQHGLV 180
 DB 121 ARPTVLATGTLVAFPTTHGDMHSLVTRHRIPEPEAAVLFROMATLAHQHGLV 180
 QY 181 RDLKLCRFVADREKKLVLENLEDSCVLTGPDSDMDKACPAVYGPETLSSRASYSK 240
 DB 157 RDLKLCRFVADREKKLVLENLEDSCVLTGPDSDMDKACPAVYGPETLSSRASYSK 205
 QY 241 AADWMSLGVALLFTMLAGHPFODSEVLLFGKIRRGAYALPAGLSAPARCLVCLRRBP 300
 DB 206 AADWMSLGVALLFTMLAGHPFODSEVLLFGKIRRGAYALPAGLSAPARCLVCLRRBP 265
 QY 301 AERLTATGILLHPWLRODMPAPTRSHLMEAAQVVDGLGDEAREEBSGDRVLYG 358
 DB 266 AERLTATGILLHPWLRODMPAPTRSHLMEAAQVVDGLGDEAREEBSGDRVLYG 323

RESULT 9
 ADC59335
 ID ADC59335 standard; protein; 349 AA.

AC ADC59335;
 DT 18-DEC-2003 (first entry)
 DE Rat cell death inhibitory protein.

XX Rat cell death inhibitory protein.

KM Rat; cell-death; cardiant; neuroprotective; anti-HIV; antiinflammatory;
 KM cerebroprotective; hepatotropic; cytosolic; immunosuppressive;
 KM antirheumatic; ophthalmological; nootropic; autophagosome;
 KM anticonvulsant; hypotensive; antiarteriosclerotic; haemostatic;
 KM antialcoholic; vinucleic; HIV; cardiac disease; immunological disease;
 KM neurodegenerative disease; ischaemic damage;
 KM central nervous system disorder; hepatic disorder;
 KM acute pancreatic inflammation; cancer; AIDS; autoimmune disease;
 KM rheumatism; Crohn's disease; glaucoma; Alzheimer's disease;
 KM Parkinson's disease; Huntington's disease; hypertension;
 KM arteriosclerosis; reperfusion damage; myocardial infarction;
 KM cerebral trauma; cerebral infarction; cerebral haemorrhage; hepatitis;
 KM alcoholic hepatitis; cerebral ischemia.

XX Rattus sp.
 XX

JP2003063986-A.
 05-MAR-2003:
 27-AUG-2001; 2001JP-00255811.
 27-AUG-2001; 2001JP-00255811.
 (SUMU) SUMITOMO SEIYAKU KK.
 WPI: 2003-601360/57.
 N-PSDB; ADCS9336.
 Controlling cell-death by administering positive or negative regulator of cell-death inhibitory-factor, for treating cancer, AIDS, autoimmune diseases, Crohn's diseases, glaucoma, Alzheimer disease.
 Claim 1; SEQ ID NO 1; 35pp; Japanese.

This invention relates to a novel method for controlling cell-death comprising administering a positive or negative regulator of cell-death inhibitory-factor. The invention also comprises a method for screening modulators of cell-death, by contacting cells expressing cell-death inhibitory factor with a candidate compound, monitoring level of expression of cell-death inhibitory factor, evaluating level of modulation ability of the compound based on change in the level of expression of the factor and selecting compounds having cell-death modulation ability. The cell death regulator of the invention may have cardiant, neuroprotective, anti-HIV, anti-inflammatory, cerebroprotective, hepatotropic, cytotactic, immunosuppressive, antineumatic, ophthalmological, nootropic, antiparkinsonia, anticonvulsant, hypotensive, antiarteriosclerotic, haemostatic, antialcoholic and vincine activities. The method of the invention is useful for treating HIV, cardiac diseases, immunological diseases, neurodegenerative disease, ischemic damage and congestion, disorder of central nervous system, hepatic disorder, acute pancreatic inflammation, and cancer, AIDS, autoimmune diseases, rheumatism, Crohn's diseases, glaucoma, Alzheimer's disease, Parkinson's disease, Huntington's disease, hypertension, arteriosclerosis, reperfusion damage, myocardial infarction, cerebral trauma, cerebral infarction, cerebral haemorrhage, hepatitis, alcoholic hepatitis, and cerebral ischemia. The present sequence represents the rat cell death inhibitory protein of the invention.

Sequence 349 AA:

Query Match 70.0%; Score 1324.5; DB 7; Length 349;
 Best Local Similarity 72.9%; Pred. No. 1.6e-120;
 Matches 261; Conservative 28; Mismatches 60; Indels 9; Gaps 3;

1 MRATPIAAPAGSLSRKKRLLELDNLTDFVQKRRSGPQRLPLPCLPLSPPTAPDRAT 60
 1 MRATSIASADVPCRKKPLFDNIDECVLRVDEPERGPTPLSP-----PPASDLSP 55
 61 AVATASRLGPIVLLPEBEGRAYOALHCPGTCTCYTVQALVLEAYALPPIKRY 120
 56 AVAPARIRLGIYLLLEBEGQNCYRALHCPGTCTCYTVQASPAQVLAAYALPPIH 115
 121 ARTEVLVAGTOLLVAFPTRTGHMNSIVRTRHRIPEBAVLFROMATALAHCHQGLV 180
 116 ARTEVLVAGTOLLVAFPTRTGHMNSIVRTRHRIPEBAVLFROMATALAHCHQGLV 175
 181 RDLKLCRFVADREKRLVLENDSCVLTPGDDSLMDKACAPAYVPELISRSASYSK 240
 176 RDLKLCRFVADREKRLVLENDSCVLTPGDDSLMDKACAPAYVPELISRSASYSK 235
 241 AADVMSIGVALFTMLAGHYPPDSEPVLLFGKTRRGAYALPAGLSAPARCLVCLRRP 300
 236 AADVMSIGVALFTMLAGHYPPDSEPVLLFGKTRRGAYALPAGLSAPARCLVCLRRP 295
 301 ABRLLTATGILLHPALRODPMPIAFTRSHLWEAAQVVDGGLDAREEGBGREVLYG 358
 296 SERLVAALGILLHPALRODPMPIAFTRSHLWEAAQVVDGGLDAREEGBGREVLYG 349

RESULT 10
 AAB43657
 ID AAB43657 standard; protein, 233 AA.

AAB43657;
 08-FEB-2001 (first entry)
 Human cancer associated protein sequence SEQ ID NO:1102.

Human; cancer associated gene; cancer antigen; detection; cancer; diagnosis; cytosolic; proliferative; vulnerability; immunomodulator; antidiabetic; antiaschmatic; antineumatic; antiarthritic; antiviral; antineoplastic; antiproliferative; antiallergic; antibacterial; cardiant; dermatologic; neuroprotective; thrombolytic; coagulant; nootropic; vasotropic; antipsoriatic; antineoplastic; gene therapy; inflammation; immune disorder; haematopoietic cell disorder; autoimmune disorder; allergic reaction; graft versus host disease; organ rejection; haemostatic; thrombolytic; cardiovascular disorder; infection; neurological disease; drug screening.

Homo sapiens.
 WO20005350-A1.
 21-SEP-2000.
 08-MAR-2000; 2000WO-US005882.
 12-MAR-1999; 99US-0124270P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM;
 WPI, 2000-587533/55.
 N-PSDB; AAC7866.

Novel isolated nucleic acids comprising sequences encoding peptides useful for treating or diagnosing e.g. cancer.

Claim 11; Page 1707-1708; 2352pp; English.

AAC78607 to AAC78448 encode the human cancer associated proteins given in AAB43398 to AAB44239. The proteins can have activities based on the tissues and cells the genes are expressed in. Example of activities include: cytosolic; proliferative; vulnerability; immunomodulator; antidiabetic; antiaschmatic; antineumatic; antiarthritic; antineoplastic; antiproliferative; antiallergic; antibacterial; antiviral; dermatologic; neuroprotective; cardiant; thrombolytic; coagulant; nootropic; vasotropic; antipsoriatic and antiangiogenic. The polynucleotides and polypeptides can be used for preventing, treating or ameliorating medical conditions and diagnosing pathological conditions. Polynucleotides, polypeptides, antibodies, agonists and antagonists from the present invention may be used to treat immune disorders by activating or inhibiting the proliferation, differentiation or mobilisation of immune cells, to treat disorders of haematopoietic cells, autoimmune disorders, allergic reactions, graft versus host disease and organ rejection, modulate haemostatic or thrombolytic activity, modulate inflammation, cancers, cardiovascular disorders, neurological disease and bacterial or viral infections. The peptides, nucleotides, antibodies, agonists and antagonists may be also be used in drug screens. AAC78449 to AAC78457 and AAB44240 represent sequences used in the exemplification of the present invention

Sequence 233 AA:

Query Match 58.2%; Score 1102; DB 3; Length 233;
 Best Local Similarity 100.0%; Pred. No. 5.1e-99;
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

151 RHRIPPEAVALFRQMAIALAHCHQGLVLRDLKLCRFVADREKRLVLENDSCVL 210

Db 26 RHRIPBEBAALVFRQMTALAHCHQHLVLDLKLCPVDFADRRKKVLLENLDSCLT 85
 Qy 211 GPDOSLWMDKACPCPYVGPBELLSSRASYSKADWWSGVAFMTLGHYPPODEPVLTF 270
 Db 86 GPDOSLWMDKACPCPYVGPBELLSSRASYSKADWWSGVAFMTLGHYPPODEPVLTF 145
 Qy 271 GKIRGAVPALPAGLSAPARCLVRCILRRPAPERTLATGILHPMLRQDPMPLAFTRSHLW 330
 Db 146 GKIRGAVPALPAGLSAPARCLVRCILRRPAPERTLATGILHPMLRQDPMPLAFTRSHLW 205
 Qy 331 EAAQVVPDGLGLDEAREEGREYVLYG 358
 Db 206 EAAQVVPDGLGLDEAREEGREYVLYG 233

RESULT 11
 ABB06093
 ID ABB06093 standard; protein; 278 AA.
 AC ABB06093;
 DT 10-MAY-2002 (first entry)
 DE Human NS protein sequence SEQ ID NO:185.
 XX
 KW Human; cytostatic; osteopathic; gynaecological; neuroprotective;
 KW antineumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;
 KW vasotropic; antiarteriosclerotic; antiinflammatory; dermatological;
 KW anorectic; muscular; antinfertility; cardiovascular; anticonvulsant;
 KW anticonvulsant; antidiabetic; tranquilliser; immunomodulator; cardiant;
 KW gastrointestinal; virucide; antitumor; antidepressant; neuroleptic;
 KW contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
 KW endometriosis; degenerative disease; multiple sclerosis; psoriasis;
 KW rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
 KW inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
 KW infertility; cardiovascular disease; coagulation disease; hypertension;
 KW ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
 KW diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
 KW gastric ulcer; Alzheimer's disease.
 OS Homo sapiens.
 XX
 PN WO200206315-A2.
 XX
 PD 24-JAN-2002.
 XX
 PF 17-JUL-2001; 2001MO-IL000653.
 XX
 PR 18-JUL-2000; 2000IL-00137345.
 XX
 PR 15-DEC-2000; 2000IL-00140354.
 XX
 PA (COMP-) COMPUGEN LTD.
 XX
 PI Mintz L, Freilich S, Bernstein U;
 XX
 DR WPI; 2002-155037/20.
 XX
 DR N-PSDB; ABR39747.
 XX
 PT One hundred and twenty eight novel nucleic acid sequences, useful for
 PT treating and diagnosing e.g. cancer, asthma and Alzheimer's.
 XX
 PS Claim 6; Page 213-214; 290p; English.
 XX
 CC ABR39691 to ABR39818 represent novel human nucleic acid sequences
 CC encoding the proteins given in ABB06037 to ABB06164. The novel sequences
 CC (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
 CC antirheumatic, antiarthritic, antipsoriatic, ophthalmological, virucide,
 CC vasotropic, antiarteriosclerotic, antiinflammatory, dermatological,
 CC anorectic, muscular, antinfertility, cardiovascular,
 CC anticonvulsant, antidiabetic, tranquilliser, immunomodulator, cardiant,
 CC immunomodulator, anticonvulsant, antidiabetic, tranquilliser, antitumor,

CC antidepressant, gastrointestinal, neuroleptic, cerebroprotective,
 CC neurotropic and contraceptive activities. The NS can be used in vaccines,
 CC gene therapy and antisense therapy. Nucleic acids, expression vectors and
 CC antibodies from the present invention can be used for treating and
 CC diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
 CC diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
 CC cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
 CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
 CC disease, coagulation disease, ischaemia, hypertension, asthma, immune
 CC disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
 CC depression, schizophrenia, viral disease, gastric ulcers, stroke,
 CC Alzheimer's disease and as a contraceptive
 XX
 SO Sequence 278 AA;

Query Match 58.0%; Score 1098; DB 5; Length 278;
 Best Local Similarity 99.1%; Pred. No. 1.ee-98;
 Matches 209; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MRATPLAAPAGSUSRRKRLDNDLTERPVOKRARGPPRLPCLLPSPAPPRAT 60
 Db 29 MRATPLAAPAGSUSRRKRLDNDLTERPVOKRARGPPRLPCLLPSPAPPRAT 88
 Qy 61 AVATASRLGPYVLLPEEGGRAYQALHCPGTGYTCVKVYVOBALAVLEPYARLPKHV 120
 Db 89 AVATASRLGPYVLLPEEGGRAYQALHCPGTGYTCVKVYVOBALAVLEPYARLPKHV 148
 Qy 121 ARTEVLVAGTQLLYAPFTTHGMSIVTRRRIPBEBAALVFRQMTALAHCHQHLV 180
 Db 149 ARTEVLVAGTQLLYAPFTTHGMSIVTRRRIPBEBAALVFRQMTALAHCHQHLV 208
 Qy 181 RDLKLCRFVADRRKKVLLENLDSCLT 211
 Db 209 RDLKLCRFVADRRKKVLLENLDSCLT 239

RESULT 12
 ABB06108
 ID ABB06108 standard; protein; 278 AA.
 AC ABB06108;
 DT 10-MAY-2002 (first entry)
 DE Human NS protein sequence SEQ ID NO:200.
 XX
 KW Human; cytostatic; osteopathic; gynaecological; neuroprotective;
 KW antirheumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;
 KW vasotropic; antiarteriosclerotic; antiinflammatory; dermatological;
 KW anorectic; muscular; antinfertility; cardiovascular; anticonvulsant;
 KW anticonvulsant; antidiabetic; tranquilliser; immunomodulator; cardiant;
 KW gastrointestinal; virucide; antitumor; antidepressant; neuroleptic;
 KW contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
 KW endometriosis; degenerative disease; multiple sclerosis; psoriasis;
 KW rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
 KW inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
 KW infertility; cardiovascular disease; coagulation disease; hypertension;
 KW ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
 KW diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
 KW gastric ulcer; Alzheimer's disease.
 OS Homo sapiens.
 XX
 PN WO200206315-A2.
 XX
 PD 24-JAN-2002.
 XX
 PF 17-JUL-2001; 2001MO-IL000653.
 XX
 PR 18-JUL-2000; 2000IL-00137345.
 XX
 PR 15-DEC-2000; 2000IL-00140354.
 XX

PA (COMP-) COMPUGEN LTD.
XX
XX Minter L, Freilich S, Bernstein J;
XX
XX MPI: 2002-155037/20.
DR N-PSDB; ABI39762.
PT One hundred and twenty eight novel nucleic acid sequences, useful for
PT treating and diagnosing e.g. cancer, asthma and Alzheimer's.
XX
XX
PS Claim 6; Page 231-232; 290pp; English.
XX
XX ABI39691 to ABI39818 represent novel human nucleic acid sequences
CC encoding the proteins given in ABB06037 to ABB06164. The novel sequences
CC (NS) can have cytostatic, osteopathic, gynecological, neuroprotective,
CC antirheumatic, antiarthritic, antipsoriatic, ophthalmological, virocidic,
CC vasotropic, antihypertensive, anti-inflammatory, dermatological,
CC anorectic, muscular, anti-HIV, antifertility, cardiovascular,
CC anticonvulsant, antifibrinolytic, hypotension, antidiabetic, tranquilizer, antiulcer,
CC immunomodulator, anticonvulsant, antidiabetic, cerebroprotective,
CC neuroprotectant, gastrointestinal, neuroleptic, cerebroprotective,
CC nociceptive and contraceptive activities. The NS can be used in vaccines,
CC gene therapy and antisense therapy. Nucleic acids, expression vectors and
CC antibodies from the present invention can be used for treating and
CC diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
CC diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
CC cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
CC disease, coagulation disease, ischemia, hypertension, asthma, immune
CC disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
CC depression, schizophrenia, viral disease, gastric ulcers, stroke,
CC Alzheimer's disease and as a contraceptive
XX
SQ Sequence 278 AA;
Query Match 58.0%; Score 1098; DB 5; Length 278;
Best Local Similarity 99.4%; Pred. No. 1.6e-98;
Matches 209; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MRATPLAAPAGSISRRKKRLDNLDTFRPVQKARSGPOPLPCLLPSPPTADPRA 60
DB 29 MRATPLAAPAGSISRRKKRLDNLDTFRPVQKARSGPOPLPCLLPSPPTADPRA 88
QY 61 AVATASRLGPVYLLEPEBEGRAYQALHCPGTETCTCKVYPVQALAVLEPYALPPHKY 120
DB 89 AVATASRLGPVYLLEPEBEGRAYQALHCPGTETCTCKVYPVQALAVLEPYALPPHKY 148
QY 121 ARPTVLVAGTQLLYAFTRTHGDMHSLVTRHRIPEPEAAVLFRQMTALAHCHQGLV 180
DB 149 ARPTVLVAGTQLLYAFTRTHGDMHSLVSPHRIPEPEAAVLFRQMTALAHCHQGLV 208
QY 181 RDLKLCRFVADREKKLVLENLEDSCVLTG 211
DB 209 RDLKLCRFVADREKKLVLENLEDSCVLTG 239
RESULT 13
ADD14087
ID ADD14087 standard; protein; 343 AA.
XX
XX ADD14087;
AC
XX
XX 01-JAN-2004 (first entry)
DT
XX
XX Human src biomarker polypeptide SEQ ID NO:276.
XX
XX
XX predictor set; protein tyrosine kinase activity modulator;
KM protein tyrosine kinase pathway; protein tyrosine kinase; cytostatic;
KM gene therapy; drug sensitivity; genetic profile; cancer; human.
XX
OS Homo sapiens.
XX
XX W02003062395-A2.

XX
XX 31-JUL-2003.
XX
XX
XX 17-JAN-2003; 2003WO-US001981.
XX
XX
XX 18-JAN-2002; 2002US-0350061P.
XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
XX Huang F, Fairchild CR, Lee FY, Shaw F,
XX
XX MPI: 2003-636735/60.
XX
XX N-PSDB; ADD14683.
XX
XX
XX New polynucleotides and polypeptides for predicting the activity of
XX PT compounds that interact with protein tyrosine kinases and/or protein
XX PT tyrosine kinase pathways.
XX
XX
XX Claim 10; SEQ ID NO 276; 139pp; English.
XX
XX
XX The present invention describes a predictor set comprising a plurality of
XX CC polynucleotides or polypeptides whose expression pattern is predictive of
XX CC the response of cells to treatment with a compound that modulates protein
XX CC tyrosine kinase activity or members of the protein tyrosine kinase
XX CC pathway. Also described: (1) predicting whether a compound is capable of
XX CC modulating the activity of cells, comprising obtaining a sample of cells,
XX CC determining whether the cells express a plurality of markers, and
XX CC correlating the expression of the markers to the compound's ability to
XX CC modulate the activity of the cells; (2) a plurality of cell lines for
XX CC identifying polynucleotides and polypeptides whose expression levels
XX CC correlate with compound sensitivity or resistance of cells associated
XX CC with a disease state; and (3) identifying polynucleotides and
XX CC polypeptides that predict compound sensitivity or resistance of cells
XX CC associated with a disease state, comprising subjecting the plurality of
XX CC cell lines to one or more compounds, analysing the expression pattern of
XX CC a microarray of polynucleotides or polypeptides, and selecting
XX CC polynucleotides or polypeptides that predict the sensitivity or
XX CC resistance of cells associated with a disease state by using the
XX CC expression pattern of the microarray. The polynucleotides and
XX CC polypeptides have cytostatic activities, and can be used in gene therapy.
XX CC The polynucleotides and polypeptides are useful in predicting the
XX CC activity of compounds that interact with protein tyrosine kinases and/or
XX CC protein tyrosine kinase pathways. These may be used in determining drug
XX CC sensitivity in patients to allow the development of individualized
XX CC genetic profiles which aid in treating diseases and disorders (e.g.
XX CC cancer) based on patient response at a molecular level. The present
XX CC sequence is used in the exemplification of the present invention.
XX
SQ Sequence 343 AA;
Query Match 39.8%; Score 753.5; DB 7; Length 343;
Best Local Similarity 47.3%; Pred. No. 9.9e-65;
Matches 167; Conservative 50; Mismatches 115; Indels 21; Gaps 6;
QY 2 RAPTPLAAPAGSISRRKKRLDNLDTFRPVQKARSGPOPLPCLLPSPPTADPRA 61
DB 5 RAPTPLAAPAGSISRRKKRLDNLDTFRPVQKARSGPOPLPCLLPSPPTADPRA 57
QY 62 VATAASRLGPVYLLEPEBEGRAYQALHCPGTETCTCKVYPVQALAVLEPYALPPHK 118
DB 58 VATAASRLGPVYLLEPEBEGRAYQALHCPGTETCTCKVYPVQALAVLEPYALPPHK 111
QY 119 HVARPTVLVAGTQLLYAFTRTHGDMHSLVTRHRIPEPEAAVLFRQMTALAHCHQGL 178
DB 112 HVARPTVLVAGTQLLYAFTRTHGDMHSLVTRHRIPEPEAAVLFRQMTALAHCHQGL 171
QY 179 VARDKLCRFVADREKKLVLENLEDSCVLTGPDLSLWPKHACPVAVGSEIISPRASY 238
DB 172 VARDKLCRFVADREKKLVLENLEDSCVLTGPDLSLWPKHACPVAVGSEIISPRASY 231
QY 239 GKADVWSLGSVALFTMLAGHYPPQDSEPVLLFGKIRRGAYALPAGLSAPARCTVRCLLR 238
DB 232 GKADVWSLGSVALFTMLAGHYPPQDSEPVLLFGKIRRGAYALPAGLSAPARCTVRCLLR 291

QY 299 EPAERLTATGILHHPMLRQDPMPLAPTRSHLMEAAQVVPDGLDEAREEGD 351
 DB 292 EFSRLTSQETLHPMFWSTD-----FVSNSAYAKAVESDQVLPDVMNEEND 339

RESULT 14
 ABB80975
 ID ABB80975 standard; protein, 372 AA.
 AC ABB80975;
 XX
 DT 21-OCT-2002 (first entry)
 XX
 DE Human tribbles homologue-1 (htrb-1) polypeptide.
 XX
 KW Human; tribbles; htrb-1; stress kinase inhibitor protein; SKIP-1; Ap-1;
 KW antithrombotic; antithrombotic; antidiabetic; antiproliferative; osteopontin;
 KW ophthalmological; cardiac; cytosolic; haemostatic; immunosuppressive;
 KW antiinflammatory; estrogen receptor; fibroblast growth factor; FGF;
 KW tumour necrosis factor; TNF.
 XX
 OS Homo sapiens.
 XX
 PN WO200253743-A2.
 XX
 PD 11-JUL-2002.
 XX
 PF 08-JAN-2002; 2002WO-US000070.
 XX
 PR 08-JAN-2001; 2001US-0260294P.
 XX
 PA (INTE-) INTERLEUKIN GENETICS INC.
 XX
 PI Dower S, Qunastrom E, Kiss-Toth E;
 XX
 DR WPI; 2002-590635/63.
 DR N-PSDB; ABN6478.
 XX
 PT Novel isolated human tribbles homologue-1 polypeptide for inhibiting Ap-1
 PT mediated inflammatory signal in a cell, and activating AP-1
 PT signal e.g. Ap-1-mediated gene activation signal in a cell.
 PS Claim 21; Fig 10B; 131p; English.
 XX

The invention provides an isolated human tribbles homologue-1 (htrb-1, also known as stress kinase inhibitor protein (SKIP-1)) polypeptide. The htrb polypeptide is useful for inhibiting an AP-1 mediated inflammatory signal in a cell. The polypeptide employed in the method is preferably htrb-1, htrb-1 N htrb-1 C, htrb-1 N C, htrb-3, htrb-3 N htrb-3 C, or htrb-3 N C. It is also useful for providing htrb agonist activity for activating an ERK-mediated signal e.g. AP-1-mediated gene activation signal, an estrogen receptor-mediated gene activation signal, an fibroblast growth factor (FGF) induced signal, or a PMA induced signal, in a cell. Htrb modulators are useful for modulating AP-1 mediated inflammatory signal in a cell such as tumor necrosis factor (TNF) induced inflammatory signal, or an interleukin induced inflammatory signal. htrb proteins are useful in screening assays, predictive medicine and in therapeutics or prophylactics. The htrb proteins are useful for screening htrb activity, such as rheumatoid arthritis, diabetes, psoriasis, osteoporosis, diabetic retinopathy, myocardial infarction and cancers. The htrb therapeutics are useful for antagonizing interleukin-1 dependent disorders of human placenta, intraventricular hemorrhage, neonatal white matter damage and subsequent cerebral palsy, and inflammation or autoimmune disorders. The present sequence represents the htrb-1 polypeptide

Sequence 372 AA;
 SQ
 XX
 XX

Query Match 39.0%; Score 737.5; DB 5; Length 372;
 Best Local Similarity 46.2%; Pred. No. 4,1e-63;
 Matches 160; Conservative 49; Mismatches 126; Indels 11; Gaps 6;

QY 2 RATPLAAPA-GSISRRKRLIED--NIDTERPVQKRRSGP--QPRLPCLPLSPPTA 55
 DB 17 RGPALLPFARVGPAPKALLPADDAVAACCPRLSECSPPDYUSPPGSPC-SPPGPPAA 75

QY 56 PDRATAVATA---SRIGPVYLLBEEGGRAVQALHCPGTGETYCKYPVQALAVLEPPYA 112
 DB 76 PGAGGSGSAPGFSRIADYLLPLAERHVSRLCHTGRRLRCKVPIKHYQDKIRPYI 135

QY 113 RLPPKHVAPPTVEYLAQTOLLYAEFTRHGDMHSLVTRHRIPEPEAAVLFQWATALAH 172
 DB 136 QLPFSHNSITGIVEYLLGETKAYVFEKSPGDMHVSVAIRKRLREEAARLFQVSAVAH 195

QY 173 CHQHGVLRLPLCKRPFARERKVLVLENDSCVLTGPPDSCIMDYACPAYGPIILS 232
 DB 196 CHQSAIVLGDCLKRKFFVSTEEERTQLTESLIEDTHIKMGDDALSDKHGCPAYVSPRIIN 255

QY 233 SRASVSGKADVMSLGVALLFTMLAGHYPPQDSEPVLLFGKIRGAVALLPAGLSAPRCIV 292
 DB 256 TTGTYSKADVMSLGVMLYTLVGRYPFHDSPPSALFSKIRKQFCIPHHISPKARCLI 315

QY 293 RCLIRRPAPERLTATGILHHPMLRQDPMPLAPTRSHLMEAAQVVPD 338
 DB 316 RSLIRREPSERLTAPETLLHPMFESVLEP-GYIDSEIGTSDQIYPE 360

RESULT 15
 AAU28135
 ID AAU28135 standard; protein, 269 AA.
 XX
 AC AAU28135;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Novel human secretory protein, Seq ID No 304.
 XX

Human; secreted protein; arthritis; Crohn's disease; sepsis; shock; ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy; transgenic animal; Alzheimer's disease; Parkinson's disease; burn; amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia; ulcer; osteoporosis; bone degenerative disorder; periodontal disease; gut protection; lung; liver fibrosis; immune deficiency; infection; severe combined immunodeficiency; SCID; autoimmune disorder; allergy; multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma; fertility; analgesic; pain; antigen.

Homo sapiens.
 WO200166689-A2.
 13-SEP-2001.
 05-MAR-2001; 2001WO-US004942.
 07-MAR-2000; 2000US-00519705.
 19-MAY-2000; 2000US-00574454.
 17-JUN-2000; 2000US-00596193.
 14-JUL-2000; 2000US-00616847.
 19-SEP-2000; 2000US-00665363.
 20-OCT-2000; 2000US-00693267.
 (HYSE-) HYSEQ INC.
 Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P, Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
 N-PSDB; AAS45035.
 WPI; 2001-589934/66.
 Novel polypeptides and nucleic acids obtained from cDNA libraries prepared from various human tissues, for diagnosis and treatment of cancer, neurological, inflammatory, and autoimmune disorders.

PS Example 4; SEQ ID NO 304; 107pp; English.

XX The invention relates to novel isolated human secreted polypeptides (I) and polynucleotides (II). (I) and (II) are useful for treating inflammatory conditions such as arthritis, nephritis, Crohn's disease, ischaemia-reperfusion injury, shock, sepsis, immune responses, and is involved in increasing haematopoiesis, stem cell survival, bone growth and remodeling. (I), (II) and modulators of (II) are useful for prophylaxis or treatment of one or more cancers. (II) is also useful for creating transgenic animals useful for studying the in vivo activities of the polypeptide as well as for studying modulators of the polypeptides. (I) induces the proliferation of neural cells and regeneration of nerve and brain tissue and is useful for the treatment of central and peripheral nervous system diseases and neuropathies, such as Alzheimer's, Parkinson's disease, Huntington's disease, and amyotrophic lateral sclerosis. In addition, (II) is involved in chemotactic or chemokinetic activity, regulation of haematopoiesis and is useful for treating myeloid or lymphoid cell disorders, platelet disorders such as thrombocytopenia and for regeneration of bone, cartilage, tendon, ligament and/or nerve tissue growth, and in tissue repair, healing of burns, incisions, ulcers, for treating osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease. Furthermore, (I) is also useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, various immune deficiencies and disorders including severe combined immunodeficiency (SCID), bacterial or fungal infections, autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic reactions and conditions, such as asthma or other respiratory problems. In addition, (I) affects biorhythms or circadian cycles of rhythms, fertility, metabolism, catabolism, anabolism, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides analgesic effects or other pain reducing effects, immunoglobulin like activity and can act as an antigen in a vaccine composition to raise an immune response. AAU28020-AAU28395 represent novel human secreted protein amino acid sequences of the invention

CC
XX
SQ Sequence 269 AA;

Query Match 35.9%; Score 678.5; DB 4; Length 269;
Best Local Similarity 51.4%; Pred. No. 1.5e-57;
Matches 132; Conservative 39; Mismatches 79; Indels 7; Gaps 2;

QY 88 CP-----TGETYCKVYVQEMALAVLEPYARLPKHVARPTVELAGTQLVAFPTRTH 141
DB 2 CPERCASTLGRRCVCFPIKHYQDKIRPYTQLPSHSNITGIWEVILGETKAYVFEKDF 61
QY 142 GDMHSLVTRTHRIPEPPAAVLPRQMATALAHQHGIVLRDILKCRFVFPADRRKKLVLE 201
DB 62 GDMHSLVTRTHRIPEPPAAVLPRQMATALAHQHGIVLRDILKCRFVFPADRRKKLVLE 121
QY 202 NLEDSCVLGGPDSIMDKHCPAYVGPBILSSRASYSKADYVSLGVALFTMLAGHYPF 261
DB 122 SLBDTHIMKGEDALSDKSGCPANVSPELINTGTYSKADYVSLGVALFTMLAGHYPF 181
QY 262 QDSEPVLLFGKIRGAYALPAGLSAPARCLVRCILRRPARELTATGILHPWLRODMP 321
DB 182 HSDPSALFSGKIRGQCFCEHISPKARCLIRSLRREPSERLTAPETILHPWFEVLEP 241
QY 322 LAPTRSHLMEAAQVVPD 338
DB 242 -GYIDSEIGTSDQIVPR 257

Search completed: August 24, 2004, 18:52:13
Job time : 63 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 18:51:12 ; Search time 18 Seconds
(without alignments)
1026.783 Million cell updates/sec

Title: US-10-070-337-5
Perfect score: 1892
Sequence: 1 MRATPLAAPAGSLSRKKRLK.....GLGLDEARBEEDGREVLYG 358

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: Issued Patents AA:*
2: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfillset1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1888	99.8	358	4	US-09-799-875-8
2	1676	88.6	360	4	US-09-509-802A-9
3	291.5	15.4	607	2	US-08-878-989-15
4	291.5	15.4	607	3	US-09-272-796-15
5	273.5	14.5	633	3	US-08-557-006C-43
6	272.5	14.4	668	4	US-09-930-181-2
7	267.5	14.1	252	4	US-07-857-224B-26
8	264.5	14.0	359	3	US-08-688-988-33
9	262	13.8	257	2	US-07-857-224B-25
10	261.5	13.8	264	2	US-07-857-224B-19
11	259.5	13.7	552	3	US-08-557-006C-40
12	257	13.6	504	4	US-09-554-726A-10
13	253	13.4	354	3	US-08-688-988-29
14	249.5	13.2	433	2	US-08-913-050A-7
15	249.5	13.2	433	2	US-08-749-902-5
16	249	13.2	264	2	US-07-857-224B-18
17	248.5	13.1	685	2	US-08-878-989-1
18	248.5	13.1	685	3	US-09-136-282-2
19	248.5	13.1	685	3	US-09-272-796-1
20	248.5	13.1	685	3	US-09-505-744-2
21	247.5	13.0	382	3	US-09-142-551A-2
22	246.5	13.0	353	3	US-08-688-988-31
23	246.5	13.0	504	4	US-09-554-726A-14
24	245	12.9	359	3	US-08-688-988-32
25	243.5	12.9	358	4	US-09-230-896C-29
26	243.5	12.9	556	4	US-09-800-960-4
27	243.5	12.9	556	4	US-10-096-960-4

ALIGNMENTS

28	243.5	12.9	565	4	US-09-800-960-2	Sequence 2, Appli
29	243.5	12.9	565	4	US-10-096-960-2	Sequence 2, Appli
30	243.5	12.9	604	4	US-09-523-849-35	Sequence 35, Appl
31	242.5	12.8	603	4	US-09-930-181-17	Sequence 17, Appl
32	241.5	12.8	351	3	US-08-688-988-28	Sequence 28, Appl
33	241	12.7	295	1	US-07-951-715A-23	Sequence 23, Appl
34	241	12.7	295	2	US-08-459-448A-23	Sequence 23, Appl
35	241	12.7	295	3	US-08-459-595A-23	Sequence 23, Appl
36	241	12.7	295	3	US-08-459-504B-23	Sequence 23, Appl
37	241	12.7	295	3	US-08-459-444-23	Sequence 23, Appl
38	241	12.7	295	4	US-09-547-422-23	Sequence 23, Appl
39	241	12.7	455	4	US-09-554-726A-7	Sequence 7, Appli
40	241	12.7	484	4	US-09-554-726A-5	Sequence 5, Appli
41	239	12.6	345	3	US-09-101-146-1	Sequence 1, Appli
42	238	12.6	273	1	US-08-252-995D-10	Sequence 10, Appl
43	238	12.6	273	2	US-08-834-108-10	Sequence 10, Appl
44	238	12.6	416	1	US-08-252-995D-2	Sequence 2, Appli
45	238	12.6	416	2	US-08-834-108-2	Sequence 2, Appli

RESULT 1
US-09-799-875-8
Sequence 8, Application US/09799875
Patent No. 6638721
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel
APPLICANT: Williamson, Mark
TITLE OF INVENTION: No. 6638721el Human Protein Kinases and Uses
FILE REFERENCE: 35800/209996
CURRENT APPLICATION NUMBER: US/09/799, 875
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: 09/659, 287
PRIOR FILING DATE: 2000-09-12
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 8
LENGTH: 358
TYPE: PRT
ORGANISM: Homo sapiens
US-09-799-875-8

Query Match 99.8%; Score 1888; DB 4; Length 358;
Best Local Similarity 99.7%; Pred. No. 8e-186; 0; Indels 0; Gaps 0;
Matches 357; Conservative 1; Mismatches 0

QY	1	MRATPLAAPAGSLSRKKRLKLELDNDLTERPVQRKASGPPRPCLPLSPPTAPDRAT	60
DB	1	MRATPLAAPAGSLSRKKRLKLELDNDLTERPVQRKASGPPRPCLPLSPPTAPDRAT	60
QY	61	AVATASRLGQVYLLEPEEGRAYQALHCPGTETTCVYVQALVLEPYARLPKHY	120
DB	61	AVATASRLGQVYLLEPEEGRAYQALHCPGTETTCVYVQALVLEPYARLPKHY	120
QY	121	ARTEVLVAGQQLYAFRTTHGMHSIVRRHILPEEAVALFRQATLAHQHGLVL	180
DB	121	ARTEVLVAGQQLYAFRTTHGMHSIVRRHILPEEAVALFRQATLAHQHGLVL	180
QY	181	RDJLCLCFVADBERKKLVLENLEDSCVLTPGPDLSLMDKACPAVYGPETILSSRASYSGK	240
DB	181	RDJLCLCFVADBERKKLVLENLEDSCVLTPGPDLSLMDKACPAVYGPETILSSRASYSGK	240
QY	241	AAVWSIGVALFTMLAGHYPPDSEPYLLFGKTRRGAYALPAGISAPARCLVRLTLREP	300
DB	241	AAVWSIGVALFTMLAGHYPPDSEPYLLFGKTRRGAYALPAGISAPARCLVRLTLREP	300
QY	301	AERLTATGILLHPRLRQDPMLAPTRSHLWEAQQVVDGGLDEARBEEDGREVLYG	358

Db 301 AERLTATGILHPMLRQDPMPLAPTRSHLMEAAQVVPDGLDREAREEGREVLYIG 358

RESULT 2
US-09-509-902A-9
Sequence 9, Application US/09509902A
Patent No. 6387676
GENERAL INFORMATION:
APPLICANT: Virca, Duke
APPLICANT: Bird, Timothy A.
APPLICANT: Anderson, Dirk M.
APPLICANT: Marken, John S.
TITLE OF INVENTION: Human cDNAs Encoding Polypeptides Having Kinase Functions
FILE REFERENCE: 2877-US
CURRENT APPLICATION NUMBER: US/09/509, 902A
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 360
TYPE: PRT
ORGANISM: Homo sapiens
US-09-509-902A-9

Query Match
Best Local Similarity 88.6%; Score 1676; DB 4; Length 360;
Matches 317; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRATPLAAPAGSISRKRRLDNDLTERPVOKRASGPPRLPCLLPSPPTAPDRAT 60
DB 43 MRATPLAAPAGSISRKRRLDNDLTERPVOKRASGPPRLPCLLPSPPTAPDRAT 102
QY 61 AVATASRLGYYVLEPEEGRAYOALHCPGTCTYTCVYVQALVLEYAALPPKHV 120
DB 103 AVATASRLGYYVLEPEEGRAYOALHCPGTCTYTCVYVQALVLEYAALPPKHV 162
QY 121 ARPEVLATGQLLYAFRTTHGDMHSIVTRHRIPEPEAAVLFROMATLALHCHQGLV 180
DB 163 ARPEVLATGQLLYAFRTTHGDMHSIVTRHRIPEPEAAVLFROMATLALHCHQGLV 222
QY 181 RDLKTRFVADDERKKLVLENLEDSCVLTGPDPSLMDKACPAVYGPETLSSRASYSCK 240
DB 223 RDLKTRFVADDERKKLVLENLEDSCVLTGPDPSLMDKACPAVYGPETLSSRASYSCK 282
QY 241 AADWMSLGVALLFTMLAGHYFQDSEPVLLPGKTRRGAYALPALSAPARCLVCLRRRP 300
DB 283 AADWMSLGVALLFTMLAGHYFQDSEPVLLPGKTRRGAYALPALSAPARCLVCLRRRP 342
QY 301 AERLTATGILHPMLRQD 318
DB 343 AERLTATGILHPMLRQD 360

RESULT 3
US-08-878-989-15
Sequence 15, Application US/08878989
Patent No. 5885803
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl G.
APPLICANT: Lal, Preeti
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
TITLE OF INVENTION: KINASES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto

STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/878,989
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0321 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 607 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1827450
US-08-878-989-15

Query Match
Best Local Similarity 15.4%; Score 291.5; DB 2; Length 607;
Matches 92; Conservative 45; Mismatches 129; Indels 49; Gaps 12;

QY 48 LPLSP-----TAPDRATAVATASRLGYPVLEPEEGG--RAYOALHCPGTCTYTCVYV 101
DB 5 LPTSDPGRLLTPRSGRTYLYKGRLLG-----KGFARCYEATDTETGSAVAIVPIQ 56
QY 102 QEALAVLEPYALP-----PKHVARPEVLAQTQLLYAFTR--THGDMHSIVR 149
DB 57 SR---VAKHOREKILNIELETRDLOHRIVRSHPEDADNIYILELCSRSKSLAHIMK 113
QY 150 TRHRIPEPEAAVLFROMATLALHCHQGLVLRDLKTRFVADDERKKLVLENLEDSCVL 209
DB 114 ARHTLLEPEVRYLYKQILSLGKLYHQRGILHRDLKGNFTT--ENNELKVGDFGLAARL 171
QY 210 TGPDSLMDKAC--PAYVGPETLSSRASYSCKAADWMSLGVALLFTMLAGHYFQDSEPV 267
DB 172 EPPEQR--KKTICGTPNYVAPVLLRQG--HGPEADWMSLGCVMYTLTGSPPEFETADLK 227
QY 268 LTFGKIRGAYALPGLSAPARCLVCLRRRPAEELTATGILHPMLRQ----DPMPLA 323
DB 228 EYRCIKQVHTLTPASLSLPARQLAATIRASPRDPSIDQLRHDFFTKGYTPDRLPS 287
QY 324 PTRSHLMEAAQVVPD 338
DB 288 -----SCVTVPD 294

RESULT 4
US-09-272-796-15
Sequence 15, Application US/09272796
Patent No. 6207148
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl G.
APPLICANT: Lal, Preeti

	APPLICANT:	Goli, Surya K.	
	TITLE OF INVENTION:	DISEASE ASSOCIATED PROTEIN	
	TITLE OF INVENTION:	KINASES	
	NUMBER OF SEQUENCES:	21	
	CORRESPONDENCE ADDRESS:		
	ADDRESSER:	Inocyte Pharmaceuticals, Inc.	
	STREET:	3174 Porter Drive	
	CITY:	Palo Alto	
	STATE:	CA	
	COUNTRY:	USA	
	ZIP:	94304	
	COMPUTER READABLE FORM:		
	MEDIUM TYPE:	Diskette	
	OPERATING SYSTEM:	IBM Compatible	
	SOFTWARE:	Fasteq for Windows Version 2.0	
	CURRENT APPLICATION DATA:		
	APPLICATION NUMBER:	US/09/272,796	
	FILING DATE:		
	CLASSIFICATION:		
	PRIOR APPLICATION DATA:		
	APPLICATION NUMBER:	08/878,989	
	FILING DATE:		
	ATTORNEY/AGENT INFORMATION:		
	NAME:	Billings, Lucy J U	
	REGISTRATION NUMBER:	36,749	
	REFERENCE/DOCKET NUMBER:	PF-0321 US	
	TELECOMMUNICATION INFORMATION:		
	TELEPHONE:	415-855-0555	
	TELEFAX:	415-845-4166	
	TELEX:		
	INFORMATION FOR SEQ ID NO:	15:	
	SEQUENCE CHARACTERISTICS:		
	LENGTH:	607 amino acids	
	TYPE:	amino acid	
	STRANDEDNESS:	single	
	TOPOLOGY:	linear	
	IMMEDIATE SOURCE:		
	LIBRARY:	GenBank	
	CLONE:	1827450	
	US-09-272-796-15		
Query Match	15.4%;	Score 291.5;	DB 3; Length 607;
Best Local Similarity	29.2%;	Pred. No. 3e-21;	
Matches	92;	Conservative 45;	Mismatches 129; Indels 49; Gaps 12
QY	48 LPISPP-----TAPDAAIVAVATSRIGPYVLLPEPESG--RAYQAALHCPGTGYTCVVYP	101	
Db	5 LPTSDGRRLITDPBSGRYYLKGRLLG-----KGGPARCYEADTDTEGSAYAKVLIPO	56	
QY	102 QEALVLEPEYARLP-----PHKIVARPTVELAGTOLLVAFFTR-THQDMNSLYR	149	
Db	57 SR---VAKXHOEKILINETELHRDIQRHLIIVAFSHHFEDADNIYIFLELSCKSKSLAHIKW	113	
QY	150 TTHRIRPERPAAVLFRMATATALACHQHGIVLVDLKICRFVADRREKKLVLENLEDSCYL	209	
Db	114 AHTTLIEPRVRYUQLILSGKLKHQRGIHLHDLKIGNFIT--ENMEIKVDFGLAART	171	
QY	210 TGPDOSLMYKNC--PAYVGPELISRASYSKGAADVMSLGALFTMLAGHYPCODSEPV	267	
Db	172 EPPEQR--KRTIGPNVYAPEVLNRQG--HGGEADVWSLGCVMVTLLCGSPPFETADLK	227	
QY	266 LTFGKRRAQVALPAGLSAPACLVRCILRRPEAEKLTATGILLHPWLRO---DEMPLA	323	
Db	228 EYTRCKQNYHTTPASTLSIPARQLLAAILRASPRDRPSIDQLRHDFEFKGYVPDRPLPTS	287	
QY	324 PTRSHLMEAQQVVD	338	
Db	288 -----SCVTVEP	294	

```

US-08-557-006C-43
; Sequence 43, Application US/08557006C
; Patent No. 6258547
; GENERAL INFORMATION:
; APPLICANT: Berl, Rajindar K.
; APPLICANT: Carling, David
; APPLICANT: Forde, Robert A.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
; FILE REFERENCE: NGAP/RH37588/UST
; CURRENT APPLICATION NUMBER: US/08/557,006C
; PRIOR FILING DATE: 1996-03-06
; PRIOR APPLICATION NUMBER: PCT/GB94/01093
; PRIOR FILING DATE: 1994-05-20
; PRIOR APPLICATION NUMBER: GB 9310489.1
; PRIOR FILING DATE: 1993-05-21
; PRIOR APPLICATION NUMBER: GB 9318010.7
; PRIOR FILING DATE: 1993-08-31
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 633
; TYPE: PRT
; ORGANISM: Yeast
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1) --(633)
; OTHER INFORMATION: Yeast SNF1 polypeptide
US-08-557-006C-43

Query Match          14.5%; Score 273.5; DB 3; Length 633;
Best Local Similarity 27.8%; Pred. No. 2.2e-19;
Matches 88; Conservative 53; Mismatches 138; Indels 37; Gaps 11,

QY 51 SPPADRAIVATATSRIGSPYVLEP-EEG--GRAYQALHCPGTGTEYCKYYPVQEAALAV 107
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 35 NSTLNNEKSSLIADAHANGYQIVKTLGDSFGKYKALVHTTTGGKVALKIIN-KKVLAK 93
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 108 LE-----PYARLPPHKVAVRTEVLAQTGLLYAEFTTRDHGMSHIVTRHRIPEPE 158
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 94 SMQGRIRREISLYRLNHPHILIKYDIKSKDEILIMVIEYAGNELFDYIVQRDMGSEGE 153
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 159 AAVLFROMATALAHCHQGLVLRDLKCRFPVADREKKLVLENLSDGCVLTGPDSDIWD 218
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 154 ARRFQQLISAVYCHRRKIKVHRDLKPEMLLDHNLVNXIADFGI--SNIMT---DGNFL 208
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 219 KIAAC--PAVYGPETLSSRAYSGKADVWSLGVALFTMLAGHYPPQDSEPVLLFGKIRRG 276
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 209 KTSICSPMYAADEVISGKL-VAGPEVDVWSCGVILYVMLCRRLPFDDESIPVLFKNISNG 267
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 277 AVALPAGISAPARCIVRCILAREPERLITATGILLHPMLRQD-PMPLAFTPSHLWEAAQV 335
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 268 VTTLPFPLSPGAAGLIKRMILVPLNKRISIHIMQDDWPKVLLPEYLLP----- 316
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 336 VPDGLDLEARREEGD 351
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 317 -PD---LKRPEEENE 328
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
US-09-930-181-2
; Sequence 2, Application US/09930181
; Patent No. 6455292
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 16U 101 V1
; CURRENT APPLICATION NUMBER: US/09/930,181
; CURRENT FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 668
; TYPE: PRT

```

ORGANISM: Homo sapiens
US-09-930-181-2

Query Match 14.4%; Score 272.5; DB 4; Length 668;
Best Local Similarity 27.6%; Pred. No. 3,1e-19;
Matches 92; Conservative 52; Mismatches 138; Indels 51; Gaps 14;

QY 65 ASRLGPVYL--LEPERGGRAYQALHCPTEGYTCVY-----YPOEALAVLEP 110
DB 13 AOVYGFYRLKTKGQGTGLVKGVHCVTCKVAIKIVNEKLSSEVLMKVERITAL-- 70
QY 111 YARLPKHVARPEVLAAGTQLYAFTTR-THGMSLVTRHRIPEEAVALFRQWATA 169
DB 71 --KLEHPHVLKADVENKRYLVLEHVSGLFPLYLVKGRLLTPEAKRPFQRTISA 128
QY 170 LAHQHGLVRLDKLRFVADREKRVKLVLENEDSCVLTPDDSLMDHAC--PAYVG 227
DB 129 LDCSHSISGHRDLPKNLL--DEKNNIRIADFGMASLOVG--DSLLET-SCGSPHYAC 183
QY 228 PELISSRASYSGKADVWSLGVALLFTMLAGHYPPDSEPVLLFGKIRGAYALPAGLSAP 287
DB 184 PEYVIRGE-KYDGRKADVWSCVILFALLVGALLPDDONLQLEKVRGVFNHPHPIPD 242
QY 288 ARCLVRCILRREPAERLTATGILHFWL---RQDPMPLAP-----TRSHLMEAAQVVP 338
DB 243 CQSLKRMIEVDARFLTLLEHIOKHIMYIGKNEPEPEQPIPRKVQIRS-LPSLEDIDPD 301
QY 339 -----GLG-----LDEAREEGDREVVLY 357
DB 302 VDSMSHSLGCFRDNKLLQDLLESEENQEMMY 334

RESULT 7
US-07-857-224B-26

Sequence 26, Application US/07857224B
Patent No. 5958784

GENERAL INFORMATION:

APPLICANT: Benner, Steven A.

TITLE OF INVENTION: Predicting Folded Structures of Proteins

NUMBER OF SEQUENCES: 114

CORRESPONDENCE ADDRESS:

ADDRESSEE: Steven A. Benner

STREET: Hadlaubstrasse 151

CITY: Zurich

STATE: none

COUNTRY: Switzerland

ZIP: (note: this is an international post code) CH-8092

COMPUTER READABLE FORM.

MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage

COMPUTER: Apple Macintosh

OPERATING SYSTEM: Macintosh 7.0

SOFTWARE: Microsoft Word

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/857,224B

FILING DATE: 03/25/92

CLASSIFICATION: 436

PRIOR APPLICATION DATA: none

TELECOMMUNICATION INFORMATION:

TELEPHONE: (International) 41 1 632 2830

TELEFAX: (International) 41 1 262 2437

TELEX: none

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 252

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE:

DESCRIPTION: protein

ORIGINAL SOURCE:

ORGANISM: Schizosaccharomyces pombe

FEATURE: Protein kinase; Table 8 Column 29

PUBLICATION INFORMATION:

AUTHORS:

AUTHORS: Hanks, S. K.
AUTHORS: Quintan, A. M.
AUTHORS: Hunter, T.
TITLE: The protein kinase family
JOURNAL: Science
VOLUME: 241
PAGES: 42-52
DATE: 1988
US-07-857-224B-26

Query Match 14.1%; Score 267.5; DB 2; Length 252;
Best Local Similarity 31.5%; Pred. No. 2,4e-19;
Matches 74; Conservative 44; Mismatches 108; Indels 9; Gaps 6;

QY 85 ALHCPTEGYTCVYPOEALAVLE-PYARLPKHVARPEVLAAGTQLLY-AFTTRHG 142
DB 20 AKAKTGDLAAIKIIPRIVASIGEMIMBLRHPNIRLYDWTDHQHYALLEVYPDG 79
QY 143 DMHSIVTRHRIPEEAVALFRQWATLAHQHGLVRLDKLRFVADREKRVKLVLEN 202
DB 80 ELPHYIRKGPUSEREPAHYLSQILDVAVACHFRFRHRLKLENTLIXVNEQO--IKI 136
QY 203 LEDSCVLTPDDSLMDHACPA--YVGPETISSRASYSGKADVWSLGVALLFTMLAGHY 260
DB 137 ADRGMAVTEPNDSCLNRY-CGSLHYLAPEIVSHK-PYRGAADVWSCVILYSLSNKLP 194
QY 261 PQDSEPVLLFGKIRGAYALPAGLSAPARCLVRCILRREPAERLTATGILHFWL 315
DB 195 FGQNTDVIVNKRIRHGAVDLPSSISSAODLLHRMIDVNPSTRTIPFESHFPL 249

RESULT 8

US-08-688-988-33

Sequence 33, Application US/08688988B

Patent No. 6096545

GENERAL INFORMATION:

APPLICANT: Lefebvre, Daniel D.

TITLE OF INVENTION: PHOSPHATE STRAVATION-INDUCIBLE PROTEINS

FILE REFERENCE: PPL96-03

CURRENT APPLICATION NUMBER: US/08/688,988B

CURRENT FILING DATE: 1996-07-31

NUMBER OF SEQ ID NOS: 48

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 33

LENGTH: 339

TYPE: PRT

ORGANISM: Glycine max

US-08-688-988-33

Query Match 14.0%; Score 264.5; DB 3; Length 339;
Best Local Similarity 28.7%; Pred. No. 7,6e-19;
Matches 86; Conservative 34; Mismatches 87; Indels 93; Gaps 12;

QY 117 HGHVARPEVLAAGTQLYAFTTRH-----GDMHSIVTRHRIPEEAVALFRQW 166
DB 58 HNNIRFEV-----FLTPHLYLVRYAAGGLPERICNAGRUSEDEARFFQDL 108
QY 167 ATALAHQHGHGLVRLDKL-----CRFVADREKRVKLVLENEDSCVLTPG 211
DB 109 ISGVSYCHSMQICHRDLKENTLIDGNPAPRLKICDFGS-----KSALLHSGPKSTVGT- 163
QY 212 PDDSLMDHACPAVYGPETISSRASYSGKADVWSLGVALLFTMLAGHYPPQDSEPVLLF- 270
DB 164 -----PAYIAPVVL-SRKEYDGKADVWSCVILYVMLVGAAYFEDPEDPKYFR 211
QY 271 --GKIRGAYALP--AGLSAPARCLVRCILRREPAERLTATGILHFWLRQD--MPL-- 322
DB 212 KSIKIRMSVOYALPYPYAVSVKRCRLISRIVANPAKINISEIKOHMFRNKLPRREITE 271
QY 323 -----APTRSHLMEAAQVVPDGLL-DEAREEGDREVV 354
DB 272 AERGRYEETQKQPSQSVSEIMQIIOEARTKIHHGEQA-----GTGISDVVRGDEANDEV 326

D _b	203	FMEEDGHKLYQCIKAGAYDFESPMDVTPEAKNLINOMLTTPAKRITAEALKEPVC	262
QY	317	Q 317	
D _b	263	Q 263	

RESULT 11
US-08-557-006C-40
; Sequence 40, Application US/08557006C
; Patent No. 6258547

```

? FILE INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINAS
? FILE REFERENCE: NGAP/PHM37588/UST
? CURRENT APPLICATION NUMBER: US/08/557,006C
? CURRENT FILING DATE: 1996-03-06
? PRIOR APPLICATION NUMBER: PCT/GB94/01093
? PRIOR FILING DATE: 1994-05-20
? PRIOR APPLICATION NUMBER: GB 9310489.1
? PRIOR FILING DATE: 1993-05-21
? PRIOR APPLICATION NUMBER: GB 9318010.7
? PRIOR FILING DATE: 1993-08-31
? NUMBER OF SEQ ID NOS: 44
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 40
? LENGTH: 552
? TYPE: prt
? ORGANISM: Rat
? FEATURE:
? NAME/KEY: gene
? LOCATION: (1) - (1747)
? OTHER INFORMATION: Full length cDNA sequence fragment of Human AMPK -
? OTHER INFORMATION: Fragment begins at nucleotide 24 and ends with
? OTHER INFORMATION: nucleotide 1765
? JS-08-557-006C-40

```

Query Match	13.7%	Score 259.5	DB 3	Length 552
Best Local Similarity	29.2%	Pred. No. 5e-18		
Matches 82; Conservative	43;	Mismatches 119;	Indels 37;	Gaps 9;

QY	6	RLGPPYL--LEPREGSAYQALHCPTECTEYTCCKYVQGE--ALALVEPYAR-----LP	115
Db	12	KIGHVILGDTGVGTFPGVKYIGEHOITGHKAAYKILNRÖKRSRLDVGKIKREIÖNKLFP	71
QY	116	PKHVARPTEVLA-GTÖLLYAFETRTGDMHSILVTRHRIEDEPZAAVLFROMATALACH	174
Db	72	RHPHILIKLQVISTPTDFEFWMEEYVSGGELPDYICKHGRVAVEVARRLFQIISAVDCH	131
QY	175	ÖHGVILRDLXLCRFVADREKKVL--LEN-----LEDSCVLTGDPDSLMDKACPA	224
Db	133	RLMVAHRDLKPEENVLLDQÖNAKTIADFGSLNMSMGDEFLRTSC-----GSPN	178
QY	225	YVPEILISRSRAYSCKADWLSLGAVFTMLACHYFPDSEPPVLLFGKIRRGAYLPAGL	284
Db	179	YAAPEVISGRL-YAGEEYDINSCGVILVALLCGTLFPDDEHVPILFLFKIKIGVPIIEYL	237
QY	265	SAPARCLRCILRRPEARLRTATGILLHPMLRÖD-PHPPLAP	324
Db	238	NSGIATILMHMLQVDFLRATIKÖIREHEWFKÖLPSTLFP	278

US-09-554-726A-10
 RESULT 12
 ; Sequence 10, Application US/09554726A
 ; Patent No. 6643369
 ; GENERAL INFORMATION
 ; APPLICANT: HERRMANN, Bernhard
 ; APPLICANT: KOSCHORZ, Birgit
 ; APPLICANT: KISPERT, Andreas

```

1  TITLE OF INVENTION: NUCLEIC ACIDS INVOLVED IN THE RESPONDER PHENOTYPE AND APPLICATION
2  TITLE OF INVENTION: THEROPF
3  FILE REFERENCE: 258.0009 0101
4  CURRENT APPLICATION NUMBER: US/09/554,726A
5  CURRENT FILING DATE: 2000-05-18
6  PRIOR APPLICATION NUMBER: PCT/EP 98/07395
7  PRIOR FILING DATE: 1998-11-18
8  PRIOR APPLICATION NUMBER: EP 98 10 3596.7
9  PRIOR FILING DATE: 1998-03-02
10 PRIOR APPLICATION NUMBER: EP 97 12 0190.0
11 PRIOR FILING DATE: 1997-11-18
12 NUMBER OF SEQ ID NOS: 53
13 SOFTWARE: PatentIn version 3.1
14 SEQ ID NO 10
15 LENGTH: 504
16 TYPE: PRT
17 ORGANISM: Mus musculus
18 US-09-554-726A-10

```

Query Match	13.6%;	Score 257;	DB 4;	Length 504;
Best Local Similarity	28.7%;	Pred. No. 8e-18;		
Matches	87;	Conservative	48;	Mismatches 122;
			Indels	46;
			Gaps	13

```

QY 71 YVLEB--BEGGRAYOL--HCPTG-----EYIC-KYVPOZALVLEBYAPLP 116
Db 28 YVLEETIGHGGQYVKAOLQHRULGNHVAKTIKKREYCNVISEVELIMAD----- 80
QY 117 HKHVARPEVLAGTOLLYAFTRTG-DMHSIVTRHRIPPEAAVLPROMATALAHQ 175
Db 81 HPRITISLLOVLETKKKVYIMELCEGKSYQHIRAGVLOHEARALFKOLISMNYCHN 140
QY 176 HGLVLDLKLCPFPADREKKVLVLENTEDSCVLTPDDSLMDKHAOPY--VGPETIJS 233
Db 141 OGIVHRDLKPDN-IMWEKDGKVKIIDFGJGTWKVKQGLNF-----CGTYFFSAPVLLS 195
QY 234 RASYSKADWMSIGVALFTMLAGHYPPQDSEPVLLFEKIRRGAYALPAGISAPARCLVR 293
Db 196 -TPYDOPKIDWVTLGVLYFVMTGKIPEDACSIKKLVRIIAGKYSIPSRISAELOJLS 254
QY 294 CLLRREPARLWATGILLHPMLRQ-----CP-----MPLAB-----TRSHIMEAAOVPR 337
Db 255 LMTANPKLRPTVAWVMHWPVTESSGVFPDPCBEQPLKPKDPALIVAMGHIGFOADIE 314
QY 338 DGL 340
Db 315 DSL 317

```

```

RESULT 13
US-08-688-988-29
; Sequence 29 Application US/06688988B
; Patent No 6065545
;
; GENERAL INFORMATION:
; APPLICANT: Lefebvre, Daniel D.
; APPLICANT: Malboubi, Mohammad A.
; TITLE OF INVENTION: PHOSPHATE STAVEATION-INDUCIBLE PROTEINS
; FILE REFERENCE: PPL96-03
; CURRENT APPLICATION NUMBER: US/08/688,988B
; CURRENT FILING DATE: 1996-07-31
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 354
;
; TYPE: PRP
;
; ORGANISM: Brassica napus
;
; IS-08-688-988-29

```

[illegible]

Db 57 HPIITREKVVLTPTHTLAIMEYAGGELFERICSGRSEDEARVFFOOLISGVSYCHA 116
QY 176 HGLVLRDLKLT-----CRFVADREKKLVLENLEDSCVLTGPDSDLMDKH 220
Db 117 MOICRDLKLENTLLDGSAPRLKICDGYG---KSLLSRPSSTVGT-----162
QY 221 ACPAVGPELISRSRYSKADAVSLGVALFTMLAGHYPFODSEPVLLF---GKIRRG 276
Db 163 --PAIAPAEVL--SRREYDGKADVWSCVTLVYMLVGAYPFEDQEDPKGFRKTIQKIMAV 219
QY 277 AYALP--AGLSAPACVLCILRRPARLTAATGLLHPW--LRQDPMPLAFTRSLM---330
Db 220 QYKIPDYVHISQDCCHLSRIEVANSILKRIITAEIKKIPWELKNLPRELTETAQAYKK 279
QY 331 -----EAAQVPPD-----GLGL-----DEAREEGDREVY 355
Db 280 ENPTSPQTAETIMKIVDDAKTPPVRSISGFGMGKGDDEEEVDEFEV 331

RESULT 14

US-08-913-050A-7
Sequence 7, Application US/08913050A
Patent No. 5827726
GENERAL INFORMATION:
APPLICANT: NEZU, Jun-ichi
TITLE OF INVENTION: DNA ENCODING PROTEIN KINASE
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROADY AND NEIMARK, P.L.L.C.
STREET: 419 7th Street N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,050A
FILING DATE: 05-SEP-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 57104/1995
FILING DATE: 16-MAR-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP PCT/JP96/00660
FILING DATE: 15-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: YUN, Allen C.
REGISTRATION NUMBER: 37,971
REFERENCE/DOCKET NUMBER: NEZU=4
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 433 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-913-050A-7

Query Match 13.2%; Score 249.5; DB 2; Length 433;
Best Local Similarity 27.0%; Pred. No. 3.8e-17;
Matches 69; Conservative 43; Mismatches 129; Indels 15; Gaps 4;

QY 117 HKHVARTEVLAG--TQLLYAFTRTGDMHSLVRT--RRRIPEEAAVLFROMATATAH 172
Db 107 HKNVIGLVLDVLYNEEKQKMTWMEYCVCGQEMLDVPEKRFVQCQHGIFCQIDLEY 166
QY 173 CHQGLVLRDLKLCRFVADREKKLVLENLEDSCVLTGPDSDLMDXACPAVVGPEILS 232

Db 167 LHSQGIWKDKIPKPNLLITGCTIKISDLGVAEHLHPAADTQRTSGSRAHQPEELAN 226
QY 233 SRASYSKADAVWSLGVALLFTMLAGHYPFODSEPVLLFGKIRGAYVALPAGLSAPARCLV 292
Db 227 GLDTFSGRKVDIWSAGVTLVNTTGLVPFEGDNIYKLFENIGKSYALPGDCGFPPLSDL 286
QY 293 RCLIRREPAEKLTAATGLLHPMLAQD-----PMPLAFTRSLHMEAAQVPP--DGLG 341
Db 287 KGMLEYEPARKRSIRQIRQSHMFRKKQHPAPAPVPIPPSPDTKORWRSMVTPYLEDLHG 346
QY 342 LDEAREEGDREVLY 357
Db 347 ADEDEDLFDIEDDIT 362

RESULT 15

US-08-749-902-5
Sequence 5, Application US/087499902
Patent No. 5985635
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goli, Surya K.
APPLICANT: Hillman, Jennifer L.
TITLE OF INVENTION: NOVEL HUMAN SERINE/THREONINE
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/749,902
FILING DATE: Filed Herewith
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0150 US
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 433 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1480861
US-08-749-902-5

Query Match 13.2%; Score 249.5; DB 2; Length 433;
Best Local Similarity 27.0%; Pred. No. 3.8e-17;
Matches 69; Conservative 43; Mismatches 129; Indels 15; Gaps 4;

QY 117 HKHVARTEVLAG--TQLLYAFTRTGDMHSLVRT--RRRIPEEAAVLFROMATATAH 172
Db 107 HKNVIGLVLDVLYNEEKQKMTWMEYCVCGQEMLDVPEKRFVQCQHGIFCQIDLEY 166
QY 173 CHQGLVLRDLKLCRFVADREKKLVLENLEDSCVLTGPDSDLMDXACPAVVGPEILS 232

```

Db      167 LHSQGIVHKDIKPGNLLITGGTLKISDILGVAAALHPFAADTCRTSGSPAFQPEIAN 226
QY      233 SRASYSKGAADVWSLIGVALFTMLAGHYPPQDSEPVLLFGKIRRGAYALPAGLSAPACTV 292
        : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      227 GLDFFSGFKVDIWSAGVTLVITTTGLIPFEGDNITYKLFNIGKSYAIPGDCGPPLSDL 286
QY      293 RCLLRREPAERLITATGILLHPMLROD-----PMELAPTTRSHLWEAQAQVP--DGLG 341
        : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      287 KGMLEYEPAKRFSIRQIRQHSWFRKHPAPAPVPIPPSPDTKDRWRSMTVVPLYEDLHG 346
QY      342 LDEAREEGRREVLY 357
        | | : : : : : : : : : : : : : : : : : : : : : :
Db      347 ADEDEDLFDIEDDITY 362

```

Search completed: August 24, 2004, 18:54:08
 Job time : 20 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 18:53:28 ; Search time 126 Seconds

(without alignments)
892.883 Million cell updates/sec

Title: US-10-070-337-5

Perfect score: 1992

Sequence: 1 MRATPLAAPAGSLSRKKRL.....GLGDDEAREBERGVLYG 358

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1295152 seqs, 314255058 residues

Total number of hits satisfying chosen parameters: 1295152

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

1: /cgn2_6/ptodata/1/pubpa/US07_PUBCOMB.pep.*

2: /cgn2_6/ptodata/1/pubpa/PTC_NEW_PUB.pep.*

3: /cgn2_6/ptodata/1/pubpa/US06_NEW_PUB.pep.*

4: /cgn2_6/ptodata/1/pubpa/US06_PUBCOMB.pep.*

5: /cgn2_6/ptodata/1/pubpa/US07_NEW_PUB.pep.*

6: /cgn2_6/ptodata/1/pubpa/PTC_NEW_PUBCOMB.pep.*

7: /cgn2_6/ptodata/1/pubpa/US08_NEW_PUB.pep.*

8: /cgn2_6/ptodata/1/pubpa/US08_PUBCOMB.pep.*

9: /cgn2_6/ptodata/1/pubpa/US09_PUBCOMB.pep.*

10: /cgn2_6/ptodata/1/pubpa/US09C_PUBCOMB.pep.*

11: /cgn2_6/ptodata/1/pubpa/US09C_NEW_PUB.pep.*

12: /cgn2_6/ptodata/1/pubpa/US10A_PUBCOMB.pep.*

13: /cgn2_6/ptodata/1/pubpa/US10A_PUBCOMB.pep.*

14: /cgn2_6/ptodata/1/pubpa/US10B_PUBCOMB.pep.*

15: /cgn2_6/ptodata/1/pubpa/US10C_PUBCOMB.pep.*

16: /cgn2_6/ptodata/1/pubpa/US10C_NEW_PUB.pep.*

17: /cgn2_6/ptodata/1/pubpa/US60_NEW_PUB.pep.*

18: /cgn2_6/ptodata/1/pubpa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1888	99.8	358	9	US-09-799-875-8
2	1888	99.8	358	12	US-10-649-156-8
3	1884	99.6	398	12	US-10-425-114-37491
4	1884	99.6	404	12	US-10-425-114-53828
5	1859	98.3	360	16	US-10-408-765A-2189
6	1676	88.6	360	14	US-10-024-828-9
7	1102	58.2	233	9	US-09-925-301-1102
8	678.5	35.9	269	12	US-10-221-278-304
9	678.5	35.9	269	12	US-10-291-172-304
10	656.5	34.7	290	12	US-10-221-278-680
11	656.5	34.7	290	15	US-10-291-172-680
12	575.5	30.4	205	14	US-10-228-263-2
13	408	21.6	138	9	US-09-864-761-45767
14	354	18.7	153	9	US-09-925-301-1367
15	344	18.2	472	12	US-10-425-114-70164

16	341.5	18.0	443	16	US-10-437-963-179120	Sequence 179120,
17	318	16.8	461	16	US-10-437-963-174166	Sequence 174166,
18	305.5	16.1	373	14	US-10-026-021-174166	Sequence 4, Appl1
19	305	16.1	477	16	US-10-437-963-122228	Sequence 122228,
20	298.5	15.8	404	16	US-10-437-963-188082	Sequence 188082,
21	296.5	15.7	778	12	US-10-423-543-11	Sequence 11, Appl1
22	296.5	15.7	778	14	US-10-354-358-92	Sequence 92, Appl1
23	296.5	15.7	778	14	US-10-116-326-2	Sequence 2, Appl1
24	294.5	15.6	431	12	US-10-424-559-146242	Sequence 146242,
25	294	15.5	794	16	US-10-311-034-17	Sequence 17, Appl1
26	293	15.5	437	12	US-10-424-599-189469	Sequence 189469,
27	293	15.5	460	12	US-10-425-114-55117	Sequence 55117, A
28	292	15.4	641	16	US-10-437-963-188700	Sequence 188700,
29	291.5	15.4	607	14	US-09-769-970-15	Sequence 15, Appl1
30	291.5	15.4	607	14	US-10-108-580-2	Sequence 2, Appl1
31	291.5	15.4	607	14	US-10-204-041-16	Sequence 16, Appl1
32	291.5	15.4	607	16	US-10-620-052A-28	Sequence 28, Appl1
33	290	15.3	434	12	US-10-425-114-49007	Sequence 49007, A
34	287.5	15.2	477	16	US-10-437-963-181140	Sequence 181140,
35	287.5	15.2	523	12	US-10-183-687-258	Sequence 258, App
36	287	15.2	511	16	US-10-437-963-125840	Sequence 125840,
37	286.5	15.1	459	12	US-10-425-114-66776	Sequence 66776, A
38	286.5	15.1	497	12	US-10-425-114-6452	Sequence 46452, A
39	286.5	15.1	754	16	US-10-479-532-1	Sequence 1, Appl1
40	285.5	15.1	467	12	US-10-425-114-57045	Sequence 57045, A
41	284	15.0	1518	9	US-09-801-368-152	Sequence 152, App
42	284	15.0	1518	15	US-10-369-493-22243	Sequence 22243, A
43	284	15.0	1518	16	US-10-618-581-11	Sequence 11, Appl1
44	283.5	15.0	448	12	US-10-424-599-236477	Sequence 236477,
45	283	15.0	504	12	US-10-183-687-407	Sequence 407, App

ALIGNMENTS

RESULT 1	
US-09-799-875-8	
Sequence 8, Application US/09799875	
Patent No. US20020034780A1	
GENERAL INFORMATION:	
APPLICANT: Meyers, Rachel	
APPLICANT: Kapeller-Libermann, Rosana	
APPLICANT: Williamson, Mark	
TITLE OF INVENTION: No. US20020034780A1e1 Human Protein Kinases and Uses	
FILE REFERENCE: 35800/209996	
CURRENT APPLICATION NUMBER: US/09/799,875	
CURRENT FILING DATE: 2001-03-06	
PRIOR APPLICATION NUMBER: 60/182,059	
PRIOR FILING DATE: 2000-02-11	
PRIOR APPLICATION NUMBER: 09/659,287	
PRIOR FILING DATE: 2000-09-12	
NUMBER OF SEQ ID NOS: 32	
SOFTWARE: FastSeq for Windows Version 4.0	
SEQ ID NO 8	
LENGTH: 358	
TYPE: PRT	
ORGANISM: Homo sapiens	
US-09-799-875-8	
Qy	1 MRATPLAAPAGSLSRKKRLDNDLTERPVQRRASGPPRLPCLLPSPPTAPPRAT 60
Db	1 MRATPLAAPAGSLSRKKRLDNDLTERPVQRRASGPPRLPCLLPSPPTAPPRAT 60
Qy	61 AVATASRLGYVLLPEPEGGRAYQALHCPGTCTYTCVYVOALAVLREYARLPKRV 120
Db	61 AVATASRLGYVLLPEPEGGRAYQALHCPGTCTYTCVYVOALAVLREYARLPKRV 120
Qy	121 ARPEVLAVGQQLVAFPTRTGDMHSIVRRIRIPEPEAAVLFRQVATALAHCHQIGVL 180

```
Db 121 ARPTEVLAGTQLLYAFRTTHGDMHSLVRSRRIPEPEAAVLFRQWATLAHCHQGLVL 180
QY 181 RDLKLCRFVADREKRLVLENLSDSCVLTGPDDSLMDKACPAVYGPETILSSRASYSK 240
Db 181 RDLKLCRFVADREKRLVLENLSDSCVLTGPDDSLMDKACPAVYGPETILSSRASYSK 240
QY 241 AADVMSLGVALFTMLAGHYPFQDSEPVLLFGKIRGAVALPAGLSAPARCLVRCILRRP 300
Db 241 AADVMSLGVALFTMLAGHYPFQDSEPVLLFGKIRGAVALPAGLSAPARCLVRCILRRP 300
QY 301 AERLTATGILHLPWLRODPMPLAPTRSHLWEAAQVVPDGLGDEAREEGDREVLYG 358
Db 301 AERLTATGILHLPWLRODPMPLAPTRSHLWEAAQVVPDGLGDEAREEGDREVLYG 358
```

```
RESULT 2
US-10-649-156-8
; Sequence 8, Application US/10649156
; Publication No. US20040038346A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Kapeller-Liebermann, Rosana
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: No. US20040038346A1e1 Human Protein Kinases and Uses
; FILE REFERENCE: 35800/209996
; CURRENT APPLICATION NUMBER: US/10/649,156
; PRIOR FILING DATE: 2003-08-27
; PRIOR APPLICATION NUMBER: US/09/799,875
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: 60/182,059
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/659,287
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 8
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-649-156-8
```

```
Query Match 99.8%; Score 1888; DB 12; Length 358;
Best Local Similarity 99.7%; Pred. No. 7.7e-159;
Matches 357; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRATPLAAGSLSRKKRLLELDNLDTERPVQKRAASGQPRLPCLPLSPPTADDRAT 60
Db 1 MRATPLAAGSLSRKKRLLELDNLDTERPVQKRAASGQPRLPCLPLSPPTADDRAT 60
QY 61 AVATASRLGPPVYLLEBEEGGRAYOALHCPTEYTKVYPVOBALAVLEPYARLPKHAY 120
Db 61 AVATASRLGPPVYLLEBEEGGRAYOALHCPTEYTKVYPVOBALAVLEPYARLPKHAY 120
QY 121 ARPTEVLAGTQLLYAFRTTHGDMHSLVTRRRIPEPEAAVLFRQWATLAHCHQGLVL 180
Db 121 ARPTEVLAGTQLLYAFRTTHGDMHSLVTRRRIPEPEAAVLFRQWATLAHCHQGLVL 180
QY 181 RDLKLCRFVADREKRLVLENLSDSCVLTGPDDSLMDKACPAVYGPETILSSRASYSK 240
Db 181 RDLKLCRFVADREKRLVLENLSDSCVLTGPDDSLMDKACPAVYGPETILSSRASYSK 240
QY 241 AADVMSLGVALFTMLAGHYPFQDSEPVLLFGKIRGAVALPAGLSAPARCLVRCILRRP 300
Db 241 AADVMSLGVALFTMLAGHYPFQDSEPVLLFGKIRGAVALPAGLSAPARCLVRCILRRP 300
QY 301 AERLTATGILHLPWLRODPMPLAPTRSHLWEAAQVVPDGLGDEAREEGDREVLYG 358
Db 301 AERLTATGILHLPWLRODPMPLAPTRSHLWEAAQVVPDGLGDEAREEGDREVLYG 358
```

RESULT 3

```
US-10-425-114-37491
; Sequence 37491, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; PRIOR FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 37491
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: Clone ID: LIB3101-194-B6_Flt_p6p
US-10-425-114-37491
```

```
Query Match 99.6%; Score 1884; DB 12; Length 398;
Best Local Similarity 99.4%; Pred. No. 2e-158;
Matches 356; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRATPLAAGSLSRKKRLLELDNLDTERPVQKRAASGQPRLPCLPLSPPTADDRAT 60
Db 41 MRATPLAAGSLSRKKRLLELDNLDTERPVQKRAASGQPRLPCLPLSPPTADDRAT 100
QY 61 AVATASRLGPPVYLLEBEEGGRAYOALHCPTEYTKVYPVOBALAVLEPYARLPKHAY 120
Db 101 AVATASRLGPPVYLLEBEEGGRAYOALHCPTEYTKVYPVOBALAVLEPYARLPKHAY 160
QY 121 ARPTEVLAGTQLLYAFRTTHGDMHSLVTRRRIPEPEAAVLFRQWATLAHCHQGLVL 180
Db 161 ARPTEVLAGTQLLYAFRTTHGDMHSLVRSRRIPEPEAAVLFRQWATLAHCHQGLVL 220
QY 181 RDLKLCRFVADREKRLVLENLSDSCVLTGPDDSLMDKACPAVYGPETILSSRASYSK 240
Db 221 RDLKLCRFVADREKRLVLENLSDSCVLTGPDDSLMDKACPAVYGPETILSSRASYSK 280
QY 241 AADVMSLGVALFTMLAGHYPFQDSEPVLLFGKIRGAVALPAGLSAPARCLVRCILRRP 300
Db 281 AADVMSLGVALFTMLAGHYPFQDSEPVLLFGKIRGAVALPAGLSAPARCLVRCILRRP 340
QY 301 AERLTATGILHLPWLRODPMPLAPTRSHLWEAAQVVPDGLGDEAREEGDREVLYG 358
Db 341 AERLTATGILHLPWLRODPMPLAPTRSHLWEAAQVVPDGLGDEAREEGDREVLYG 398
```

```
RESULT 4
US-10-425-114-53828
; Sequence 53828, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; PRIOR FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 53828
; LENGTH: 404
; TYPE: PRT
```

ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Clone ID: LIB082-008-C10_Flt pep
US-10-425-114-53828

Query Match 99.6%; Score 1884; DB 12; Length 404;
Best Local Similarity 99.4%; Pred. No. 2e-158;
Matches 356; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MRATPLAAPAGSLSRKRLLELDNDLTERPVOKRARGSPQRLPCLPLSPPTADPDRAT 60
DB 47 MRATPLAAPAGSLSRKRLLELDNDLTERPVOKRARGSPQRLPCLPLSPPTADPDRAT 106
QY 61 AVATASRLGPPVLLPEBEGGRAVOALHCPGTGTEYCKVYPVOEALAVLEPYARLPKHV 120
DB 107 AVATASRLGPPVLLPEBEGGRAVYRALHCPGTGTEYCKVYPVOEALAVLEPYARLPKHV 166
QY 121 ARPEVLATGTLVAFPTRTGDMHSLVTRHRIPEPEAAVLFRQMATLALHCHQGLVL 180
DB 167 ARPEVLATGTLVAFPTRTGDMHSLVSRHRIPEPEAAVLFRQMATLALHCHQGLVL 226
QY 181 RDLKLCRFVADREKKVLLENLSDSCVLTGPDSDLMDKACPAAYVGPETLSSRASYSKG 240
DB 227 RDLKLCRFVADREKKVLLENLSDSCVLTGPDSDLMDKACPAAYVGPETLSSRASYSKG 286
QY 241 AADVMSLGVALFTMLAGHYPFQDSEPVLLFGKIRGAYALPAGLSAPARCLVRLRRP 300
DB 287 AADVMSLGVALFTMLAGHYPFQDSEPVLLFGKIRGAYALPAGLSAPARCLVRLRRP 346
QY 301 AERLTATGTLHPMLRQDMPPLAPTRSHLMEAAQVVDGLGIDEAREEGDREVVLYG 358
DB 347 AERLTATGTLHPMLRQDMPPLAPTRSHLMEAAQVVDGLGIDEAREEGDREVVLYG 404
```

RESULT 5

US-10-408-765A-2189
Sequence 2189, Application US/10408765A
Publication No. US20040101874A1
GENERAL INFORMATION:
APPLICANT: Ghosh, Soumitra S.
APPLICANT: Fahy, Bojin D.
APPLICANT: Zhang, Bing
APPLICANT: Gibson, Bradford W.
APPLICANT: Taylor, Steven W.
APPLICANT: Glenn, Gary M.
APPLICANT: Marnock, Dale E.
TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
FILE REFERENCE: 660088.465
CURRENT APPLICATION NUMBER: US/10/408,765A
CURRENT FILING DATE: 2003-04-04
NUMBER OF SEQ ID NOS: 3077
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2189
LENGTH: 360
TYPE: PRF
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: 193, 194, 195, 196, 197
OTHER INFORMATION: Xaa = Any Amino Acid
US-10-408-765A-2189

Query Match 98.3%; Score 1859; DB 16; Length 360;
Best Local Similarity 98.3%; Pred. No. 2.9e-156;
Matches 354; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

```
QY 1 MRATPLAAPAGSLSRKRLLELDNDLTERPVOKRARGSPQRLPCLPLSPPTADPDRAT 60
DB 1 MRATPLAAPAGSLSRKRLLELDNDLTERPVOKRARGSPQRLPCLPLSPPTADPDRAT 60
QY 61 AVATASRLGPPVLLPEBEGGRAVOALHCPGTGTEYCKVYPVOEALAVLEPYARLPKHV 120
```

```
DB 61 AVATASRLGPPVLLPEBEGGRAVOALHCPGTGTEYCKVYPVOEALAVLEPYARLPKHV 120
QY 121 ARPEVLATGTLVAFPTRTGDMHSLVTRHRIPEPEAAVLFRQMATLALHCHQGLVL 180
DB 121 ARPEVLATGTLVAFPTRTGDMHSLVSRHRIPEPEAAVLFRQMATLALHCHQGLVL 180
QY 181 RDLKLCRFVADREKKVLLENLSDSCVLTGPDSDLMDKACPAAYVGPETLSSRASYS 238
DB 181 RDLKLCRFVADREKKVLLENLSDSCVLTGPDSDLMDKACPAAYVGPETLSSRASYS 240
QY 239 GKADVMSLGVALFTMLAGHYPFQDSEPVLLFGKIRGAYALPAGLSAPARCLVRLRR 298
DB 241 GKADVMSLGVALFTMLAGHYPFQDSEPVLLFGKIRGAYALPAGLSAPARCLVRLRR 300
QY 299 EPARLTATGTLHPMLRQDMPPLAPTRSHLMEAAQVVDGLGIDEAREEGDREVVLYG 358
DB 301 EPARLTATGTLHPMLRQDMPPLAPTRSHLMEAAQVVDGLGIDEAREEGDREVVLYG 360
```

RESULT 6

US-10-024-828-9
Sequence 9, Application US/10024828
Publication No. US20030036051A1
GENERAL INFORMATION:
APPLICANT: Vircs, Duke
APPLICANT: Bird, Timothy A.
APPLICANT: Anderson, Dirk M.
APPLICANT: Marken, John S.
TITLE OF INVENTION: Human CDNA's Encoding Polypeptides Having Kinase
TITLE OF INVENTION: Functions
FILE REFERENCE: 2877-US
CURRENT APPLICATION NUMBER: US/10/024,828
CURRENT FILING DATE: 2001-12-18
PRIOR APPLICATION NUMBER: US/09/509,902A
PRIOR FILING DATE: 1999-08-03
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 360
TYPE: PRF
ORGANISM: Homo sapiens
US-10-024-828-9

Query Match 88.6%; Score 1676; DB 14; Length 360;
Best Local Similarity 99.7%; Pred. No. 4.9e-140;
Matches 317; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MRATPLAAPAGSLSRKRLLELDNDLTERPVOKRARGSPQRLPCLPLSPPTADPDRAT 60
DB 43 MRATPLAAPAGSLSRKRLLELDNDLTERPVOKRARGSPQRLPCLPLSPPTADPDRAT 102
QY 61 AVATASRLGPPVLLPEBEGGRAVOALHCPGTGTEYCKVYPVOEALAVLEPYARLPKHV 120
DB 103 AVATASRLGPPVLLPEBEGGRAVOALHCPGTGTEYCKVYPVOEALAVLEPYARLPKHV 162
QY 121 ARPEVLATGTLVAFPTRTGDMHSLVTRHRIPEPEAAVLFRQMATLALHCHQGLVL 180
DB 163 ARPEVLATGTLVAFPTRTGDMHSLVSRHRIPEPEAAVLFRQMATLALHCHQGLVL 222
QY 181 RDLKLCRFVADREKKVLLENLSDSCVLTGPDSDLMDKACPAAYVGPETLSSRASYSKG 240
DB 223 RDLKLCRFVADREKKVLLENLSDSCVLTGPDSDLMDKACPAAYVGPETLSSRASYSKG 282
QY 241 AADVMSLGVALFTMLAGHYPFQDSEPVLLFGKIRGAYALPAGLSAPARCLVRLRRP 300
DB 283 AADVMSLGVALFTMLAGHYPFQDSEPVLLFGKIRGAYALPAGLSAPARCLVRLRRP 342
QY 301 AERLTATGTLHPMLRQD 318
DB 343 AERLTATGTLHPMLRQD 360
```

RESULT 7

US-09-925-301-1102
; Sequence 1102, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1102
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1102

Query Match 58.2%; Score 1102; DB 9; Length 233;
Best Local Similarity 100.0%; Pred. No. 2,2e-89;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 RHRIPPEAAVLFROMATLALHCHQGLVLRDLKLCRFVADREKRLVLENTEDSCVLT 210
DB 26 RHRIPPEAAVLFROMATLALHCHQGLVLRDLKLCRFVADREKRLVLENTEDSCVLT 85
QY 211 GPDDSLMDKACPAVYGPETLSRASYSKGAADVSLGVALFTMLAGHPFQDSEPVLLF 270
DB 86 GPDDSLMDKACPAVYGPETLSRASYSKGAADVSLGVALFTMLAGHPFQDSEPVLLF 145
QY 271 GKIRGAYALPAGLSAPARCLVRCILRREPARELTATGILHFWLMDPMPPLAPTESHLM 330
DB 146 GKIRGAYALPAGLSAPARCLVRCILRREPARELTATGILHFWLMDPMPPLAPTESHLM 205
QY 331 EAAQVVPDGLGDEAREEEDREVLYG 358
DB 206 EAAQVVPDGLGDEAREEEDREVLYG 233

RESULT 8
US-10-221-278-304
; Sequence 304, Application US/10221278
; Publication No. US20040034208A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: No. US20040034208A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: US/10/221,278
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 304
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-221-278-304

Query Match 35.9%; Score 678.5; DB 12; Length 269;
Best Local Similarity 51.4%; Pred. No. 9.4e-52;

Matches 132; Conservative 39; Mismatches 79; Indels 7; Gaps 2;
QY 88 CP-----TGTETCKYYPVGEALAVLEPYARLPKHKVAPTEVLATQLLVAFRTTH 141
DB 2 CPGRCASTLGRVYCKVFPIKHVODKIRPYIQLPESHNITGIVILLGETATVYFFEKDF 61
QY 142 GDMHSIVTRHRIPEPEAAVLFROMATLALHCHQGLVLRDLKLCRFVADREKRLVLE 201
DB 62 GDMHSIVTRHRIPEPEAAVLFROMATLALHCHQGLVLRDLKLCRFVADREKRLVLE 121
QY 202 NLEDSCVLTGPDDSLMDKACPAVYGPETLSRASYSKGAADVSLGVALFTMLAGHPF 261
DB 122 SLEDTHIMKCEDDALSKHGCPCAVYSEIINTGTYSKGAADVSLGVALFTMLAGHPF 181
QY 262 QDSEPVLLFGKIRGAYALPAGLSAPARCLVRCILRREPARELTATGILHFWLMDPMP 321
DB 182 HSDPSALFSGKIRRGQCEPIEHISPAKCLIRSLRREPARELTATGILHFWLMDPMP 241
QY 322 LAPTRSHLMEAAQVVD 338
DB 242 -GYTDSIGTSDQIVPE 257

RESULT 9
US-10-291-172-304
; Sequence 304, Application US/10291172
; Publication No. US20030228584A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: No. US20030228584A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: US/10/291,172
; CURRENT FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 304
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-291-172-304

Query Match 35.9%; Score 678.5; DB 15; Length 269;
Best Local Similarity 51.4%; Pred. No. 9.4e-52;
Matches 132; Conservative 39; Mismatches 79; Indels 7; Gaps 2;

QY 88 CP-----TGTETCKYYPVGEALAVLEPYARLPKHKVAPTEVLATQLLVAFRTTH 141
DB 2 CPGRCASTLGRVYCKVFPIKHVODKIRPYIQLPESHNITGIVILLGETATVYFFEKDF 61
QY 142 GDMHSIVTRHRIPEPEAAVLFROMATLALHCHQGLVLRDLKLCRFVADREKRLVLE 201
DB 62 GDMHSIVTRHRIPEPEAAVLFROMATLALHCHQGLVLRDLKLCRFVADREKRLVLE 121
QY 202 NLEDSCVLTGPDDSLMDKACPAVYGPETLSRASYSKGAADVSLGVALFTMLAGHPF 261
DB 122 SLEDTHIMKCEDDALSKHGCPCAVYSEIINTGTYSKGAADVSLGVALFTMLAGHPF 181
QY 262 QDSEPVLLFGKIRGAYALPAGLSAPARCLVRCILRREPARELTATGILHFWLMDPMP 321
DB 182 HSDPSALFSGKIRRGQCEPIEHISPAKCLIRSLRREPARELTATGILHFWLMDPMP 241
QY 322 LAPTRSHLMEAAQVVD 338

QY 264 SEPVLLFGKIRGAYVALPAGLSAPARCLVRLRREPARELTATGILHPWLRODPMLA 323
Db 121 SDPSALFSKIRRGQCFEHSIPSPARCLIRSLRREPSERLTAPBILHPFESVLEP-G 179
QY 324 PTRSHLWEAQQVDPD 338
Db 180 YIDSEIGTSDQVPE 194

RESULT 13
US-09-864-761-45767
; Sequence 45767, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hamel, David R.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 45767
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009486.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.93
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.59
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.9
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.93
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.96

OTHER INFORMATION: SWISSPROT HIT: O74536, EVALUATE 1.00e-18
; OTHER INFORMATION: EST_HUMAN HIT: BE897149.1, EVALUATE 2.00e-71
US-09-864-761-45767

Query Match 21.6%; Score 408; DB 9; Length 138;
Best Local Similarity 63.1%; Pred. No. 3.8e-28;
Matches 77; Conservative 15; Mismatches 30; Indels 0; Gaps 0;

QY 197 KLVLENLSDSCVLTGPDSDIMDKHACPAVYGPETLSSRASVSGKADVMSIGVALFTMLA 256
Db 2 RVLKESIEDAVYLRGDDSDSDKHGCPAYVSPETLNTSSGSGKADVMSIGVALFTMLV 61

QY 257 GHYFQDSEPVLLFGKIRGAYVALPAGLSAPARCLVRLRREPARELTATGILHPWL 316
Db 62 GRYPFHDIPESSLFSSKIRRGQFNIPELSPKACLIIRSLRREPSERLTQELIDHPWFS 121

QY 317 QD 318
Db 122 TD 123

RESULT 14
US-09-925-301-1367
; Sequence 1367, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1367
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (136)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (138)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (141)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (142)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (143)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (152)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-925-301-1367

Query Match 18.7%; Score 354; DB 9; Length 153;
Best Local Similarity 93.3%; Pred. No. 2.7e-23;
Matches 70; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 MRATPLAAPAGSLSRKKRLIEDNLDTERPVOKARSGPQRLPPCLPLSPPTAPDRAT 60
Db 74 MRATPLAAPAGSLSRKKRLIEDNLDTERPVOKARSGPQRLPPCLPLSPPTAPDRAT 133

QY 61 AVATASRLGPVYLE 75
Db 134 AVXTXSRKXXYYVLE 148

```
RESULT 15
US-10-425-114-70164
; Sequence 70164, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 70164
; LENGTH: 472
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73099C12_F11.pep
US-10-425-114-70164

Query Match      18.2%; Score 344; DB 12; Length 472;
Best Local Similarity 33.3%; Pred. No. 9e-22; Indels 28; Gaps 10;
Matches 98; Conservative 44; Mismatches 124;

QY 51 SP-PTAPPRATVATASR--LGPVYL--LPEEGGAAVQALHCPTEYTCXYVQEA 104
Db 3 SPSPQPD---MVSQARGLGAYELGRTLGNGFGKVKQARHRSQGQFAVK--IMER 56
QY 105 IAVLEPYA-----RLPPKHVARFTEVLASTQLLYAFTRTHGD--MSIVRTR 151
Db 57 ANVLQGRADQICREIATLKLAAHPVVALHEVAASKTKIYVWLELVNNGELIDRIASE 116
QY 152 HRIPEPAALVFRQNTALAHQHGHGLVLRDLKLCRFVADRERK-KLVLENLEDSCVLT 210
Db 117 GRLPEQEARRLFQQLVDGVSCHKEGVCHRDLLK-ENVLVDKGNIKISDFGSLALPQHL 175
QY 211 GDDDSIMDMGACPAVYVGPRLSSRASYSGRADVMSIGVALFTMLAGHYPFQDSEPVLLF 270
Db 176 GNDGLHTTCGSPNTIAPVYLQNR-GYDGLSDIWSGVLIVMLVGHLPFDDRNTIVLY 234
QY 271 GKIRGAYALPAGISAPARCLVRCILRREPAPERLTATGILLHPMLRQDPMPLAP 324
Db 235 QKIFKDAQIFEWLSPGARNTLRLRIEPPDAERIAMAEIKAHMPWFQEHVVFVLP 288
```

Search completed: August 24, 2004, 19:04:28
Job time : 129 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 18:50:17 ; Search time 19 seconds
(without alignments)
1812.451 Million cell updates/sec

Title: US-10-070-337-5

Perfect score: 1892
Sequence: 1 MRLTPLAAPAGSLSRKKRL.....GLGLDEAREEGDREVLVYG 358

Scoring table: BIOSDM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	299.5	15.8	631	2 A57286	probable serine/th
2	298	15.8	887	2 T20941	hypothetical prote
3	291.5	15.4	651	2 S52244	p69g3 protein - A
4	290	15.3	602	2 S72513	FOG2 protein - yea
5	284	15.0	1518	2 S37928	probable purine nu
6	283	15.0	504	2 T10449	probable serine/th
7	283	15.0	512	2 T52633	serine/threonine-s
8	280	14.8	512	1 JCI446	serine/threonine-s
9	279	14.7	339	2 S56719	probable protein k
10	276.5	14.6	502	2 T02306	hypothetical prote
11	276	14.6	441	2 E85362	serine/threonine-s
12	276	14.6	511	1 A56009	hypothetical prote
13	274	14.5	431	2 T02496	probable protein k
14	273.5	14.5	633	1 A26030	serine/threonine-s
15	273	14.4	533	1 S60304	serine/threonine-s
16	272	14.4	504	2 T07415	probable serine/th
17	271.5	14.3	426	2 C71408	probable carbon ca
18	270	14.3	480	2 T41587	probable serine/th
19	269.5	14.2	480	2 A86427	hypothetical prote
20	269	14.2	512	2 T07788	hypothetical prote
21	268.5	14.2	421	2 E96522	Ca2+/calmodulin-de
22	268.5	14.2	533	1 A34366	protein kinase cdr
23	267.5	14.1	593	1 KIZPMN	SNF-related kinase
24	266	14.1	472	2 B90100	Ca2+/calmodulin-de
25	266	14.1	542	1 A45025	serine/threonine-s
26	264	14.0	622	1 S44859	Ca2+/calmodulin-de
27	263	13.9	542	1 A26464	Ca2+/calmodulin-de
28	263	13.9	589	2 S68470	protein kinase AK2
29	262.5	13.9	421	2 T48202	

30	262	13.8	1142	2 S59359	G1N4 protein - yea
31	261	13.8	442	2 T48203	hypothetical prote
32	260.5	13.8	513	1 S60303	serine/threonine-s
33	260.5	13.8	552	1 S51025	[hydroxymethyl]glut
34	259.5	13.7	473	1 S59941	serine/threonine-s
35	259.5	13.7	552	1 A53621	[hydroxymethyl]glut
36	259.5	13.7	891	2 T40503	protein kinase kin
37	259.5	13.7	891	2 A38903	protein kinase 1 -
38	258	13.6	350	2 T06107	probable serine/th
39	258	13.6	726	2 T33998	hypothetical prote
40	257.5	13.6	440	2 T14736	calcium-dependent
41	257.5	13.6	492	1 T03271	calcium-dependent
42	257	13.6	469	2 B84644	probable protein k
43	257	13.6	502	1 A41361	serine/threonine-s
44	255.5	13.5	445	2 T50802	serine/threonine p
45	255	13.5	610	1 A49082	calcium-dependent

ALIGNMENTS

RESULT 1

A57286
probable serine/threonine protein kinase (EC 2.7.1.-) fnk - mouse

C:Species: Mus musculus (house mouse)

C:Date: 01-Dec-1995 #sequence_revision 01-Dec-1995 #text_change 10-Sep-1997

C:Accession: A57286

R:Donohue, P.D.; Alberts, G.F.; Guo, Y.; Winkles, J.A.

J. Biol. Chem. 270, 10351-10357, 1995

A:Title: Identification by targeted differential display of an immediate early gene enc

A:Reference number: A57286; MUID:95247749; PMID:7730342

A:Accession: A57286

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-631 <DON>

A:Cross-references: GB:U21392; GB:U22434

C:Superfamily: unassigned Ser/thr or Tyr-specific protein kinases; protein kinase homol

C:Keywords: ATP; phosphotransferase

F:61-315/Domain: protein kinase homology <KIN>

Query Match

Best Local Similarity 28.3%; Pred. No. 3,8e-13; Length 631;

Matches 98; Conservative 48; Mismatches 133; Indels 67; Gaps 14;

QY	39	POPRUPCLLPSPPTAPDRATAVATAS-RLGPVYLPE-----	77
DB	10	PRP-PPRAVPSAPAPGPG--PPNAPSREPEVLACPRADPPGRLLTIDLSGRTYK 65	
QY	78	-----EGG--RAYQALHCPTGTEYCKVYPVOEALAVLEPYARLP-----PKKA 119	
DB	66	GRLLGKGPRACRYEATDTSGIAVAVKVIPOQR---VAKPHQREKLTINEIELHRDLQHRH 122	
QY	120	VAREPEVLAGTOLLYAFETR-THGDMHSLVTRRHRPEPEAAVLFPOMATLALHCHQHG 178	
DB	123	YRFSHFHPDADNIVYFELCSKRSLSAHMKARHTLLEBVRYYLLSGLKYLHORG 182	
QY	179	VLRDLKLCRPVPRDRKKLVLENLSDCVLTGPDSDMDKAC--PAVGEPEILSSRAS 236	
DB	183	LRHDLKLGNEFTTD--NMELKVGDFGLARLEPPEQR--KTTICGPNVYAPVLLRQG- 237	
QY	237	YSGKADWMSLGVALLFTMLAGHPFODSEFVLLFGKIRGAYVALPAGLSAPARCLVRCCL 296	
DB	238	-HPEADWMSLGCVMVTLTLCGSPFFETADLKEYRCIKQVHTLPASLSIPARQLLAAIL 296	
QY	297	RREPARLTVATGILLHPWLQ---DMPILAPTRSHLMEAAQVVPD 338	
DB	297	RASPBRPSPSEQLIRHDFFTKGYTPDRLPVS-----SCVTVPD 334	

RESULT 2

T20941
hypothetical protein F15A2.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
 C/Accession: T20941
 R/Gregory, J.
 Submitted to the EMBL Data Library, March 1996
 A/Reference number: Z19349
 A/Accession: T20941
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-887 <WILL>
 A/Cross-references: EMBL:Z70207; PIDN:CAA94127.1; GSPDB:GN00028; CESP:FL5A2.6
 A/Experimental source: clone FL5A2
 C/Genetics:
 A/Gene: CESP:FL5A2.6
 A/Map position: X
 A/Intons: 32/1; 63/3; 92/2; 139/2; 189/3; 328/2; 448/2; 516/3; 604/2; 684/3; 735/2; 777

Query Match 15.8%; Score 298; DB 2; Length 887;
 Best Local Similarity 25.9%; Pred. No. 7e-13;
 Matches 95; Conservative 51; Mismatches 133; Indels 88; Gaps 14;

QY 62 VATASTRIGPYVL--LEPEEGRAYQALHCPTEGYCKV-----YVQELAV 107
 DB 11 VAAQAYCGPYKLEKLGQGTGLVKGTGHCITGRKAIVKNEKLSSEVLQKVEREIAI 70
 QY 108 LEFYALRPHKHVARPTVAVAGTOLYAFTR-THGDMSLVTRRHRIPEPEAAVLFROM 166
 DB 71 M-----KLIEPHVLHYDYENKRYLYLLLEHVSGETLFDYVRKRLMSKEARKFRO 126
 QY 167 ATALAHCHQGLVRLDLKCRFVFADREKK-----LVLEN-LEDSCVITGDDSL 216
 DB 127 ISALDFCHAHNIQRLKXENLILDERNNIKAVDFGMSLOVEGSMLETSQ----- 177
 QY 217 WDKHACPAVGPBILSSRASYSGKADVSLGVALFTMLAGHPFODSEPVLLFGKIRRG 276
 DB 178 ----GSPHYACEVIRGE-KYDGRKADVMSCGVILLVALLGALPFDNDILRLILEKVKRG 232
 QY 277 AVAALPAGSAPARCIVRCILRREPARLTAATGILLHFWL-----RQDP--MPLAR-TRSH 328
 DB 233 VHIHFVFPADVQSLIRAMIEVDPGKRYSLADVFKHPWVSGETTKADPELELPMQVQTH 292
 QY 339 IWEAAQV-VPGIG-----IDENAE---EGGD 351
 DB 293 VIPGEDSIDPVLRLMNCIGCFKDKOKILNELSPRANTKRVVFLDLDRKRREPADD 352
 QY 352 REVLYYG 358
 DB 353 TEIVLRG 359

RESULT 3
 S52244
 p69g3 protein - African clawed frog
 C/Species: Xenopus laevis (African clawed frog)
 C/Date: 07-May-1995 #sequence_revision 03-Aug-1995 #text_change 24-May-2001
 C/Accession: S52244
 R/Rochi, C.; le Guellec, R.; Paris, J.; Couturier, A.; Philippe, M.
 Submitted to the EMBL Data Library, October 1992
 A/Description: E93, selected by differential screening encodes a new Xenopus protein kin
 A/Reference number: S52243
 A/Accession: S52244
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-651 <ROG>
 A/Cross-references: EMBL:Z17205; MID:9609283; PIDN:CAA78913.1; PID:9609284
 C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homo
 C/Keywords: ATP
 F/11-265/Domain: protein kinase homology <KIN>
 F/19-27/Region: protein kinase ATP-binding motif

Query Match 15.4%; Score 291.5; DB 2; Length 651;
 Best Local Similarity 29.9%; Pred. No. 1.4e-12;
 Matches 73; Conservative 47; Mismatches 103; Indels 21; Gaps 7;

QY 85 ALHCPTEGYCKVYVQELAVLEPYARLP-----PHKVARPTVLAAGTOLYAF 137
 DB 30 ASHLITGEKAIVKIMD-KESIGDLPVKTEIDAMKULSHQVCLRHVITPKKIFVYL 88
 QY 138 TR-THGDMSLVTRRHRIPEPEAVLRQMATALAHCHQGLVRLDLKCRFVADREKK 196
 DB 89 EYCPGSELFDYITAKDLTEEARVFFROIVSAVAVYHSQYARHDLKPEMLIDEDQNL 148
 QY 197 KLVLENEDSCVLTPDDSLMDKH-----ACPAVGPBILSSRASYSGKADVSLGVAL 251
 DB 149 KLI-----DPLCAKPKGGL-DYHLMTCGSPAYABELIQKA-YIGSERDIWSWGLM 201
 QY 252 FTMLAGHPPODEPVLLFKIRRGAYADAGSAPARCIVRCILRREPARLTAATGIL 311
 DB 202 YALMCGLPPDDNDVWLVYKKIMRGKYEIPKMLSPGVSLLSQMVDPKKRITVHILN 261
 QY 312 HPWL 315
 DB 262 HPWL 265

RESULT 4
 S72513
 FOG2 protein - yeast (Kluyveromyces marxianus var. lactis)
 C/Species: Kluyveromyces marxianus var. lactis; Candida sphaerica
 C/Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 07-May-1999
 C/Accession: S72513
 R/Boffin, P.; Ficarella, A.; Domini, C.; Lodi, T.; Puglisi, P.P.; Ferrero, I.
 Curr. Genet. 29, 316-326, 1996
 A/Title: FOG1 and FOG2 genes, required for the transcriptional activation of glucose-reg
 A/Reference number: S72513; MID:96171514; PMID:8598052
 A/Accession: S72513
 A/Status: not compared with conceptual translation
 A/Molecule type: DNA
 A/Residues: 1-602 <GOF>
 A/Note: the source is designated as Kluyveromyces lactis
 C/Genetics:
 A/Gene: FOG2
 C/Function:
 A/Description: probably involved in the regulation of glucose-repressible gene expressio
 C/Superfamily: AMP-activated protein kinase; protein kinase homology
 F/33-286/Domain: protein kinase homology <KIN>

Query Match 15.3%; Score 290; DB 2; Length 602;
 Best Local Similarity 28.5%; Pred. No. 1.6e-12;
 Matches 88; Conservative 51; Mismatches 138; Indels 32; Gaps 10;

QY 63 ATASRIQPYVLEP-ERG--GRAYQALHCPTEGYCKVYVQELAVLE-----P 110
 DB 27 AOGHIGKQIITKTLGSGFQKVLAVHISTGQVALKIN-KVYLAKSDMQRIEREIS 85
 QY 111 YARLPKHNARPTVAVAGTOLYAFTRTHGDMSLVTRRHRIPEPEAAVLFROMATL 170
 DB 86 YLRLLRPHIILKLDVIAKSDIIMVLEYAGNELFDYIVQDKMBQBARFPQOIIISAV 145
 QY 171 AHCHQGLVRLDLKCRFVFADREKKLVLENEDSCVLTPDDSLMDKHAC--PAYVG 228
 DB 146 DYCHRHICIVRDLKPEMLIDENHINVKIADFG--SNIMT---DGNFLKISCGSPNNAAP 200
 QY 229 EILSSRASYSKADVSLGVALFTMLAGHPFODSEPVLLFGKIRRGAYALPAGLSAPA 288
 DB 201 EVISGKL-YAGPEVDWSSGVLLYVMLCRRLPFODESIPVLFKNIISGVYTIIPNPLSGA 259
 QY 289 RCIVRCLRREPARLTAATGILHFWLROD--PMPLATRSGLWEEAQAQVVDGLGDARE 347
 DB 260 ASLTKKMLVNPVNPRIIVHIMQDEWPKVLLPVLVPAESTHQENS-----ESKT 309
 QY 348 EEDGREVYL 356
 DB 310 EDDGPSVPL 318

RESULT 5

S37928
Probable purine nucleotide-binding protein YKL101w - Yeast (Saccharomyces cerevisiae)
N/Alternate names: protein YKL453
C/Species: Saccharomyces cerevisiae
C/Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 24-Sep-1999
C/Accession: S37928; S39084
R/Cheret, G.; Fukuhara, H.; Bolotin-Fukuhara, M.; Dajman-Fornier, B.; Pallier, C.; Puzo
submitted to the Protein Sequence Database, March 1994
A/Reference number: S37920
A/Accession: S37928
A/Molecule type: DNA
A/Residues: 1-1518 <CHR>
A/Cross-references: EMBL:Z28101; NID:G486168; PIDN:CAA81941.1; PID:G486169; MIPS:YKL101w
A/Experimental source: strain S288C
R/Pallier, C.; Valens, M.; Puzos, V.; Fukuhara, H.; Cheret, G.; Sor, F.; Bolotin-Fukuhara
Yeast 9, 1149-1155, 1993
A/Title: DNA sequence analysis of a 17 kb fragment of yeast chromosome XI physically loc
protein kinases.
A/Reference number: S39084; MUID:94078677; PMID:8256524
A/Accession: S39084
A/Status: translation not shown
A/Molecule type: DNA
A/Residues: 1-1518 <PBL>
A/Cross-references: EMBL:X71133; NID:G431205; PIDN:CAA50456.1; PID:G431215
A/Experimental source: strain S288C
C/Genetics:
A/Gene: SGD:HSU1
A/Cross-references: SGD:S0001584; MIPS:YKL101w
A/Map position: 11L
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C/Keywords: ATP; P-loop; purine nucleotide binding; serine/threonine-specific protein ki
F:79-369/Domain: protein kinase homology <KIN>
F:79-86/Region: nucleotide-binding motif A (P-loop)
F:87-95/Region: protein kinase ATP-binding motif
F:85/Binding site: ATP/GTP (Lys) #status Predicted

Query Match 15.0%; Score 284; DB 2; Length 1518;
Best Local Similarity 26.4%; Pred. No. 1.2e-11;
Matches 85; Conservative 51; Mismatches 128; Indels 58; Gaps 9;

QY 56 PDRATAVATAS-----RLGPVLT--LEPEGGRAVYALCPGTETCTCYVPOEYL 105
DB 59 PDSTVAVATSSKRSKRDVGVKGLGKLGSGSVRLAKMKTGQLAAIKIVPKKAF 118
QY 106 -----AVLPYA-----RLPPEKGVAR 122
DB 119 VHCNNGTVPNSYSSMTVSNVSSPSIASREHSNHSQTPYGIERTIVIMKLSHNVVA 178
QY 123 PTEVLATGQLVAFTRTH-GDMHSLVTRRHRIPEPEAAVLFROMATALAHCHQGLVLR 181
DB 179 LFEVWENKSELIVILEVVDGELFDVLVSKGLPEREALHYFQIVGVSYCHSFICHR 238
QY 182 DLKLCRFVADREKKLVLENEDSCVLTGPDDSLMDKACPAIVYGPILLSSASTSGKA 241
DB 239 DLK-PENLLLDKNRRIKLADEF-GMAALELPNKLITKSCSPVYASPEIWMGE-PYHGGP 295
QY 242 ADVWMSGVALFTMLAGHYPPQDSEPVLLFGKIRRGAYALPAGSAPARCLVRCILRREPA 301
DB 296 SDWMSGIIVFALLTGLLPNDNITKLLKLVOSGKIOMPSNLSSEARDISKLIVDPPE 355
QY 302 ERLTATGILLHPWLRO-DPEML 322
DB 356 KRITTOELLKHPILIKKDDLPV 377

RESULT 6
T10449
Probable serine/threonine-specific protein kinase (EC 2.7.1.1) - cucumber
N/Alternate names: SNF1-related protein kinase
C/Species: Cucumis sativus (cucumber)
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 20-Jun-2000
C/Accession: T10449
R/Gumpel, N.J.

submitted to the EMBL Data Library, December 1996
A/Reference number: Z17020
A/Accession: T10449
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-504 <GUM>
A/Cross-references: EMBL:Y10036
A/Experimental source: cv. Masterpiece; coryledon
C/Function:
A/Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin
C/Superfamily: AMP-activated protein kinase; protein kinase homology
C/Keywords: AMP; phosphotransferase; serine/threonine-specific protein kinase
F:6-260/Domain: protein kinase homology <KIN>

Query Match 15.0%; Score 283; DB 2; Length 504;
Best Local Similarity 28.3%; Pred. No. 4e-12;
Matches 86; Conservative 48; Mismatches 116; Indels 54; Gaps 10;

QY 80 GRAVQALHCPTEGYTCVY-----PVOEALAVLEPYARLPKHKVAPTEVL 127
DB 20 GKVKIAEHALTGKVAIKILNRKIKULDMEKYRREIKIL----RLFMHPIILRVLEVI 75
QY 128 AGTQLYAFPTRT-HGDMHSLVTRRHRIPEPEAAVLFROMATALAHCHQGLVLRDLK-- 184
DB 76 ETPSDIYVMEYVKSGLFDYIVKGRLOEDEARNFQQLISGVYCHRNWVHRDLKPE 135
QY 185 -----LCRFVADREKKLVLEN--LEDSCVLTGPDDSLMDKACPAIVYGPILLSSASTSGKA 236
DB 136 NLLDSSKNVNIAPFGGSLNIRRDGHFLKTS-----GSPVYAAPEVLSGKL- 181
QY 237 YSGRAVDWMSGVALFTMLAGHYPPQDSEPVLLFGKIRRGAYALPAGSAPARCLVRCIL 296
DB 182 YAGPEVDWMSGVALIYALICGTLFPDDENINMLPKIKGIGIYTLPSHSSGARELIPML 241
QY 297 RREPARLTATGILLHPWLRODPMPLAPTRSHLEMAQVDPDGLDPAEEBEGD--REV 354
DB 242 VVDPMKRITLIPETIRQHWNF-----QALHPRYLAVPPEDT-MQAKKIDEDILQEV 290
QY 355 VLYG 358
DB 291 VMWG 294

RESULT 7
T52633
Serine/threonine-specific protein kinase (EC 2.7.1.1) AKIN1 [validated] - Arabidopsis
N/Alternate names: SNF1 protein kinase omolog AKIN1
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 03-Nov-2000
C/Accession: T52633
R/Bhalerao, R.P.; Salchert, K.; Bako, L.; Okresz, L.; Szabados, L.; Muranaka, T.; Mach
Proc. Natl. Acad. Sci. U.S.A. 96, 5322-7, 1999
A/Title: Regulatory interaction of PRL1 WD protein with Arabidopsis SNF1-like protein
A/Reference number: Z25116; MUID:99238528; PMID:10220464
A/Accession: T52633
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-512 <BHA>
A/Cross-references: EMBL:X99279; PIDN:CAA67671.1
A/Experimental source: cultivated Columbia
C/Genetics:
A/Gene: AKIN11
C/Function:
A/Description: EC 2.7.1.1; serine/threonine-specific protein kinase AKIN1 [validated,
complements SNF1 mutations in yeast
C/Superfamily: AMP-activated protein kinase; protein kinase homology
C/Keywords: AMP; phosphotransferase; serine/threonine-specific protein kinase

Query Match 15.0%; Score 283; DB 2; Length 512;
Best Local Similarity 28.6%; Pred. No. 4.1e-12;
Matches 87; Conservative 48; Mismatches 115; Indels 54; Gaps 10;

QY 80 GRAVQALHCPTEGYTCVY-----PVOEALAVLEPYARLPKHKVAPTEVL 127

```

Db      32 GKVIAEIVLVGHKVAIKILNRRKIKNMEMEKVRREIKIL-----RLFMHHIIRLOEVI 87
QY      128 AGTQLVAFPTRT-HGDMHSLVTRTHRIPEPEAVLPRQATATALAHCHQGLVLRDK- 184
Db      88 ETTSDIVMEVYVSGELFDYIVVEKGRLOEDEANFPQOIIISGVEYCHRNWVVRDJKPE 147
QY      185 -----LCRFVADREKKVLLEN--LEDSCVLTPDDSLMDKACAPAYVGETLSSRAS 236
Db      148 NLLDSRCNKVIADFGLSNVWRDGHFLKTSC-----GSPNYAAPEVTSGL- 193
QY      237 YSGKAADWMSLGVALLFTMLAGHYFPDSEVLLFGKIRGAVALPAGLSAPARCLVACL 296
Db      194 YAGEVDVWVSCGVLLVALLCGTLFPDDENIPNLFKTKIGGITYLPSHLSSEARDLIPRL 253
QY      297 RREPARELTATGILLHPMLRQDPMPLAPTRSHLMEAAQVVPDGLDEARE--EEGREGV 354
Db      254 IYDPVKRVTIPETIRQHPWF-----QTHLPRTYLAVPPT-VEQAKKINEEIVQEV 302
QY      355 VLYG 358
Db      303 VNWG 306

```

RESULT 8

JCI446

serine/threonine-specific protein kinase (EC 2.7.1.-) AK21 - Arabidopsis thaliana

N/Alternate names: protein kinase SNF1 homolog

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #ext_change 11-Jun-1999

C/Accession: JCI446; S58266; S66334

R/Author: T. Thomas, M.; Bianchi, M.; Halford, N.G.; Kreis, M.

Gene 120, 249-254, 1992

A/Title: Structure and expression of a gene from Arabidopsis thaliana encoding a protein

A/Reference number: JCI446; MUID:93013041; PMID:1339373

A/Accession: JCI446

A/Molecule type: DNA

A/Residues: 1-512 <LEG>

A/Cross-references: GB:M93023; NID:g166599; PIDN:AAA2736.1; PID:g166600

R/Thummler, F.; Kirchner, M.; Teuber, R.; Ditttrich, P.

submitted to the EMBL Data Library, May 1995

A/Description: Differential accumulation of the transcripts of 22 novel protein kinase

A/Reference number: S58266

A/Accession: S58266

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 144-198 <THU>

A/Cross-references: EMBL:X86966; NID:g928909; PIDN:CAA60529.1; PID:g928910

R/Thummler, F.; Kirchner, M.; Teuber, R.; Ditttrich, P.

Plant Mol. Biol. 29, 551-565, 1995

A/Title: Differential accumulation of the transcripts of 22 novel protein kinase genes

A/Reference number: S66334; MUID:96123233; PMID:8534852

A/Accession: S66334

A/Molecule type: DNA

A/Residues: 144-198 <TH2>

A/Cross-references: EMBL:X86966; NID:g928909; PIDN:CAA60529.1; PID:g928910

C/Comment: This enzyme plays an important role in a signal transduction cascade regulat

C/Genetics:

A/Gene: AK101; AK21

A/Introns: 64/1; 125/3; 186/3; 230/3; 292/3; 322/3; 350/3; 396/3; 475/3

C/Function:

A/Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin

C/Superfamily: Am-activated protein kinase; protein kinase homology

C/Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kinase

F:17-271/Domain: protein kinase homology <KIN>

F:25-33/Region: protein kinase ATP-binding motif

F:48-67,142,144/Active site: Lys, Glu, Asp, Lys #status predicted

F:147,151/Binding site: magnesium (Asn, Asp) #status predicted

```

QY      80 GRAVQALHCPGTETCTKV-----PVQALAVLEPYARLPKHVAPTEVL 127
Db      31 GRVIAEHALTGHKVAIKILNRRKIKNMEMEKVRREIKIL-----RLFMHHIIRLOEVI 86
QY      128 AGTQLVAFPTRT-HGDMHSLVTRTHRIPEPEAVLPRQATATALAHCHQGLVLRDK- 184
Db      87 ETPDITLVMEVYVSGELFDYIVVEKGRLOEDEANFPQOIIISGVEYCHRNWVVRDJKPE 146
QY      185 -----LCRFVADREKKVLLEN--LEDSCVLTPDDSLMDKACAPAYVGETLSSRAS 236
Db      147 NLLDSRCNKVIADFGLSNVWRDGHFLKTSC-----GSPNYAAPEVTSGL- 192
QY      237 YSGKAADWMSLGVALLFTMLAGHYFPDSEVLLFGKIRGAVALPAGLSAPARCLVACL 296
Db      193 YAGEVDVWVSCGVLLVALLCGTLFPDDENIPNLFKTKIGGITYLPSHLSSEARDLIPRL 252
QY      297 RREPARELTATGILLHPMLRQDPMPLAPTRSHLMEAAQVVPDGLDEAREEGDEEV 355
Db      253 VYDPMKRVITIPETIRQHPWF-----QAHLPRTYLAVPPTVOQAKKIDEEIIOEVI 302
QY      356 LYG 358
Db      303 NMG 305

```

RESULT 9

S56719

serine/threonine-specific protein kinase SPK-1 (EC 2.7.1.-) - soybean

C/Species: Glycine max (soybean)

C/Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #ext_change 18-Jun-1999

C/Accession: S56719

R/Author: Yoon, H.W.; Jeong, Y.H.; Bahk, J.D.; Hong, J.C.; Cho, M.J.

submitted to the EMBL Data Library, January 1993

A/Description: Cloning of a novel protein serine/threonine kinase cDNA from soybean.

A/Reference number: S56719

A/Accession: S56719

A/Molecule type: mRNA

A/Residues: 1-339 <SHI>

A/Cross-references: EMBL:L01453; NID:g169990; PIDN:AAA33979.1; PID:g169991

C/Superfamily: kinase-related transforming protein; protein kinase homology

C/Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase

F:3-261/Domain: protein kinase homology <KIN>

F:11-19/Region: protein kinase ATP-binding motif

Query Match

Best Local Similarity 14.7%; Score 279; DB 2; Length 339;

Matches 86; Conservative 34; Mismatches 86; Indels 88; Gaps 11;

```

QY      117 HKHVARPTREVLAGTQLVAFPTRT-----GDMHSLVTRTHRIPEPEAVLPRQ 166
Db      58 HPNIIIRFEV-----FLTPHTALVLERYAAGBELFERICNAGRUSEDEARFFQOL 108
QY      167 ATALAHCHQGLVLRDKL-----CRFVADREKKVLLENLEDSCVLTG 211
Db      109 ISGVSYCHSMGICHRDLKENTLLDGNPAPRLKICDFES---KSAIHSQPKSTVGT- 163
QY      212 PDDSLMDKACAPAYVGPILSSRASYSKADWMSLGVALLFTMLAGHYFPDSEVLLF- 270
Db      164 -----PAYIAEVL-SRKEYDGKADVMSCGVTLVMLVGAVPEPDPEPKNFR 211
QY      271 ---GKIRGAVADP--AGLSAPARCLVRLRREPARELTATGILLHPMLRQD--PML- 322
Db      212 KSIKRLMSVOYALPDYVRVSKCHNLISCIFVANPARISISEKOHMFKXNIPREITE 271
QY      323 -----APTRSHLMEAAQV-VPDGLDEAREE 348
Db      272 PERRGVVDHRTQPSQSVERTWRIIGEARIKIHGEQAGTSTDAVGDANEE 325

```

RESULT 10

T02306

probable protein kinase (imported) - Arabidopsis thaliana

N/Alternate names: hypothetical protein F13p17.2

C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001
C:Accession: J02306; D84753
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul
submitted to the EMBL Data Library, July 1998
A:Description: Arabidopsis thaliana chromosome II BAC F13P17 genomic sequence.
A:Reference number: Z14657
A:Accession: J02306
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-502 <ROT>
A:Cross-references: EMBL:AC004481; NID:G3337347; PIDN:AAC27394.1; PID:G3337349
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.L.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffett, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;
Euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: AB4420; MUID:20083487; PMID:10617197
A:Accession: D84753
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-502 <STD>
A:Cross-references: GB:AE002093; NID:G3337349; PIDN:AAC27394.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g34180, F13P17.2
A:Map position: 2
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homold
F/55-311/Domain: protein kinase homology <KIN>

```

Query Match      14.6%; Score 276.5; DB 2; Length 502;
Best Local Similarity 26.0%; Pred. No. 1..le-11;
Matches 89; Conservative 56; Mismatches 13; Indels 61; Gaps 12;

Qy 3 ATPLAP-----AGSLRKRRLLELDNIDTERPVQKARSGBQRLPCLLPSPT 54
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 6 STPLALPGPIPIQMGLLAR--IVTKNKNETSPBSPRS-----PRTP----- 48

Qy 55 AEDRAVAATAASRLGPAVLLPEPEGRAYOALHCPGTGYTCRYKP-----VQ 102
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 49 ---QGSLMDKYELTGK---LYGHSGFAKVYLAARNTHSGSDVAIKVIDKEKIYKSLAGHIK 103

Qy 103 EALAIVEPARLPENHGVARPTEVLATQTLYAFPTTTH-GDMISLVETRRIRIPEEAAV 161
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 104 REISIL----RRVRHPYIVHLIEVMATKTYYIMAEYRGSELYNTV-ARGLRGGTLAR 158

Qy 162 LFROMATALAHCHONGVLBDLKLCRFVEADREKKVLENLEDSCVLTPGDDSLMDKHA 221
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 159 YFOOLISSVAFCHRSGVYHRDLKLENLILDCKGNKYK-----SDFGLSVSEOLKOEI 212

Qy 222 C-----PAYVEPILLSRASYSKGADWMSIGVALFTMLAGHYPOSEVPILFEKIR 275
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 213 CGFFCGTPAYIAPEVL-TRKGYEGAKDIWSGVLLPYLMAGYLPPDDKNIIIVTYKIK 271

Qy 276 GAYALPAGLSAPARCIVRCILRRPEARELTATGILLHPWLKQ 317
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 272 GGFPCPKMFEPFLARLVTRMIDLNDPTRITTIPEIKKHREFKK 313

RESULT 11
EB85362
hypothetical protein AT4g30960 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #next_change 02-Mar-2001
C:Accession: EB85362
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: AB5001; MUID:20083488; PMID:10617198
A:Accession: B85362
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-441 <STO>
```

A:Cross-references:	GB:NC_001268; NID:g7269998; PIDN:CAB79814.1; GSPDB:GN00140
C:Genetics:	
A:Gene:	AT4g30960
A:Map position:	4
C:Superfamily:	unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol
Query Match	14.6%; Score 276; DB 2; Length 441;
Best Local Similarity	29.4%; Pred. No. 1e-11;
Matches	77; Conservative 38; Mismatches 125; Indels 22; Gaps 4;
QY	81 RAYQALHCPGTETCYCKY-----PVBALAVLEPYARLPKHKVARPTELYA 128
DB	37 KYVHARNIQTGKSVAMKRVGKEKVKVGMVDQIKREISVM----RMVGHPIIVLEHVEWYA 92
QY	129 GTQLLIYAFTTTHGDMHSVTRHRIPEPEAAVLFRQMTALAHQHQGVLRDLKCRF 188
DB	93 SKSKIYFAMEIYVRGSELPKAVAKGRLREDVARYVFOQLISAVDFCHSRGCVYHRDLKPENTL 152
QY	189 VFADBERKKVLVENEDSCVLTPDDSLMDHACAPAYGPEITLSRASYSGKAADWSIG 248
DB	153 LIDEGNLKVVDPGISAFPTENLKODGLDHTTCGIPAYVAPEVITLKK-GYDGKAKDLMSGC 211
QY	249 VALFTMLAGHYPEQDSEPVLLFGKIRRGAYVLPAGLSAPARCLVCLTRPEABRLITATG 308
DB	212 VILFTLLGLGYLPQDDNLVNMVTRKLYRGDFKCPGWLSSDARLVLTKLDDPNENTRIITLX 271
QY	309 ILLHFWLRQDPM-----PLAPT 325
DB	272 VMDSFWFKQATRSRNPVAAAT 293

```

RESULT 12
A56009
serine/threonine-specific protein kinase (EC 2.7.1.-) NP65 - common tobacco
C|Species: Nicotiana tabacum (common tobacco)
C|Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 16-Jun-2000
C|Accession: A56009
R:Murataka, T.; Banno, H.; Machida, Y.
Mol. Cell. Biol. 14, 2958-2965, 1994
A|Title: Characterization of tobacco protein kinase NP65, a homolog of Saccharomyces cerevisiae
A|Reference number: A56009; MUID:94217693; PMID:8164654
A|Accession: A56009
A|Status: preliminary
A|Molecule type: mRNA
A|Residues: 1-511 <MUR>
A|Cross-references: GB:D26602; NID:3496384; PIDD:BA05649.1; PID:3496385
C|Function:
A|Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonine-phosphate
C|Superfamily: AMP-activated protein kinase; protein kinase homology
C|Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kinase
F:17-271/Domain: protein kinase homology <KIN>
F:25-33/Region: protein kinase ATP-binding motif
F:48,67,142,144/Active site: Lys, Glu, Asp, Lys #status predicted
F:147,151/Binding site: magnesium (Asn, Asp) #status predicted

Query Match          14.6%; Score 276; DB 1; Length 511;
Best Local Similarity 28.3%; Pred. No.1.2e-11;
Matches 86; Conserved 49; Mismatches 115; Indels 54; Gaps 10;

Qy      80 GRAVQALHCPTEGTEVYCKV-----PVGALAVLEPYARLPKHVAPTEVL 127
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      31 GKVKIAEHLTLGHKVAVKILNRKKIKMMEEEKVREIKTL-----RLFMHPIITRIYEV 86

Qy      128 AGTQLLYAFETR-HDDMSLVTRNRIRPEPAALVFRQATALACHONGVLRLDKLC 186
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      87 ETPSDVIYVMEVYKSGLEFDYIEKGLQDDEKPKFFQDLISGVETCHRMMVNRDLPKE 146

Qy      187 RFVFADRRERKVL---LEN-----LEDSCVLTGPDDSLMDKACRAVYGPETLSSRAS 236
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      147 NLLLDGKMMVKINDPGLSINMRDGHFLKTSK-----GSPFNVAARVIGSKL- 192

Qy      237 YSGKADWVSLGVALTMTLAGHYPQDSEFVLLFGKIRGAVNLPAGLSAPACIVRCLL 296

```

```

Db      193 YAGEVWVWSCGVTITVALCTTLPEDDENTPMLFKKIKGMIISPHSLSAGARDLIPML 252
Oy      297 RREBAERLTATGILHFLRLRODPMPLAPTRSHLEAQAQVVDGGLDEAREED--REV 354
Db      253 IVDPMKMTIPEINMHEMF-----QMLPRYLAVPPTD-MQOAKKIDEDIIQEV 301
Oy      355 VLYG 358
Db      302 VKRG 305

```

RESULT 13

```

T02496
probable protein kinase [imported] - Arabidopsis thaliana
N:Alternate names: hypochemical protein T19C21.2
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001
C/Accession: T02496; F84805
R:Roundley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul
submitted to the EMBL Data Library, August 1998
A:Description: Arabidopsis thaliana chromosome II BAC T19C21 genomic sequence.
A:Reference number: Z14676
A:Accession: T02496
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-431 <ROU>
A/Cross-references: EMBL:AC004683; NID:G3395421; PID:G3395423
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Roundley, S.D.; Shee, T.P.; Bentio, M.T.; Town, C.D.; Fujii, C.Y.;
euse, D.; Nlemtan, W.C.; Cronin, L.A.; Shen, M.; Vanakem, S.B.; Umayam, L.; Tallon, L.;
Nature 402, 761-768, 1999
A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; NCID:20083487; PMID:10611197
A:Accession: F84805
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-431 <STO>
A/Cross-references: GB:AE002093; NID:G3786023; PID:AA67369.1; GSPDB:GN00139
C:Genetics:
A:Gene: T19C21.2 Atg38490
A:Map position: 2
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo

```

```

Query Match      14.5%; Score 274; DB 2; Length 431;
Best Local Similarity 29.5%; Pred. No. 1.4e-11;
Matches 92; Conservative 45; Mismatches 141; Indels 34; Gaps 9;

Oy      54 TAPDRATAVATASRLGPGVYLLEPEEGRAYOALHCPGTGE-YTCVKYP-----Y 101
Db      16 TDDDNKSAFLFGKYDLGK--LLGSAFAKYQAEIDONGESVAIKVQKKRLKDLTNAV 73
Oy      102 GEAALVLEAVYARLPKHKVAPTEVLA-GTOLLVAFRTTHGDMISLVTRTRIPPEEA 160
Db      74 KEISIMW---RLRHPIHVLSEVALATKTKYFWELAKKGELPSRV-TSNRFESISR 128
Oy      161 VLEFQWATLAHCHOGHVLRLDKLCRFVADREKRLVLEEDSCVLGTGDDSLMKH 220
Db      129 KVFRLTISAVRYCHARGVHRDLKPENLLIDNRDLKVSDFGLSMKEGIDHPDMLHTTC 188
Oy      221 ACPAVYGEIILSSRASYSKADAVSLGVALFTMLAGHYFPDSEPVLLFGKIRGAVAL 280
Db      189 GTPAVVADELILKK-CYDGSKADIMSCGVLFLLNAGYLPFRDPIMIGLYKRIHQAQKL 247
Oy      281 PAGESAPARCVRCILRREPARELTATGILHPLMLRODPMPLAPTRSHLEAQAQVVDG 340
Db      248 PDWTSDDRKLRLRLBPNRPLITVEETLKDPWF-----NHGVDPSEIT--GI 294
Oy      341 GLDEAREEEDGR 352
Db      295 QADVDYDLRENCK 306

```

RESULT 14

```

A26030
serine/threonine-specific protein kinase (EC 2.7.1.1-) SNF1 - Yeast (Saccharomyces cerev.
N:Alternate names: protein YDR477w
C:Species: Saccharomyces cerevisiae
C>Date: 20-Aug-1987 #sequence_revision 20-Aug-1987 #text_change 23-Mar-2001
C/Accession: A26030; S63644
R:Celenza, J.U.; Carlson, M.
Science 233, 1175-1180, 1986
A>Title: A yeast gene that is essential for release from glucose repression encodes a p
A:Reference number: A26030; MUID:86289463; PMID:3526554
A:Accession: A26030
A:Molecule type: DNA
A:Residues: 1-633 <CEL>
A/Cross-references: EMBL:M13971; NID:G172629; PID:AAA35058.1; PID:G172630
R:Dietrich, F.S.
submitted to the EMBL Data Library, August 1995
A:Description: The sequence of S. cerevisiae cosmid 9410, 8035, 8166, and 9787.
A:Reference number: S63654
A:Accession: S63644
A:Molecule type: DNA
A:Residues: 1-633 <DIE>
A/Cross-references: EMBL:U33050; NID:G927726; PID:AA64904.1; PID:G927732; GSPDB:GN000C
C:Genetics:
A:Gene: SGD:SNF1; MTPS:YDR477w
A/Cross-references: SGD:S0002885; MTPS:YDR477w
A:Map position: 4R
C/Function:
A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin
A>Note: required for expression of glucose-repressed genes in response to glucose depriv
C:Superfamily: AMP-activated protein kinase; protein kinase homology
C:Keywords: ATP; autophosphorylation; magnesium; nucleus; phosphoprotein; phosphotransfe
F:33-306/Domain: protein kinase ATP-binding motif
F:51-69/Region: protein kinase ATP-binding motif
F:84,103,177,179/Active site: Lys, Glu, Asp, Asn #status predicted
F:182,186/Binding site: magnesium (asn, Asp) #status predicted
F:210/Binding site: phosphate (Thr) (covalent) (by unidentified kinase) #status predicte

```

```

Query Match      14.5%; Score 273.5; DB 1; Length 633;
Best Local Similarity 27.8%; Pred. No. 2.3e-11;
Matches 88; Conservative 53; Mismatches 136; Indels 37; Gaps 11;

Oy      51 SEPTAPDRATAVATASRLGPGVYLLEP-EG--GRAYOALHCPGTGEYTCVKYPVDALAV 107
Db      35 SSTLNKPKSLADAGHIGNYQIVKTLGSGRGKVALAHTTGGQVAKLIIN-KKVLAK 93
Oy      108 LE-----PYARLPKHKVAPTEVLAGTQLLYAFRTTHGDMHSLVTRRHIPPEE 158
Db      94 SPMQGRIREREISYLLRLHPIHVLSEVALATKTKYFWELAKKGELPSRV-TSNRFESISR 153
Oy      159 AAVLFQWATLAHCHOGHVLRLDKLCRFVADREKRLVLEEDSCVLGTGDDSLMD 218
Db      154 ARRFQOQIISAVEYCHRHKIVARDLKPENLLIDENLVKIDPGL--SNIMT--DGNFL 208
Oy      219 KHAC--PAYVGEIILSSRASYSKADAVSLGVALFTMLAGHYFPDSEPVLLFGKIRRG 276
Db      209 KTCGSEPVYAPAEVLSGL-YAGEVVDWVSCGVTITVALCTTLPEDDENTPMLFKKIKGMI 267
Oy      277 AYALPAGUSAPARCVRCILRREPARELTATGILHPLMLRODPMPLAPTRSHLEAQAQV 335
Db      268 VYTLPEKFLSPGAAGIKMLVLVPLNRSIHHEIMQDDMFXYDLPYLLP----- 316
Oy      336 VPDGIGLDEAREEEDGR 351
Db      317 -PD--LKPHEEENE 328

```

RESULT 15

```

S60304
serine/threonine-specific protein kinase (EC 2.7.1.1-) BKIN12 (version 2) - barley
C:Species: Hordeum vulgare (barley)
C>Date: 19-Mar-1997 #sequence_revision 15-Aug-1997 #text_change 11-Jun-1999

```

C:Accession: S60304; S24579
R:Halford, N.G.; Vicente-Carbajosa, J.; Sabelli, P.A.; Shewry, P.R.; Hannappel, U.; Kretz Plant J. 2, 791-797, 1992
A:Title: Molecular analyses of a barley multigene family homologous to the yeast protein
A:Reference number: S60303; MUID:93258420; PMID:1302632
A:Accession: S60304
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-513 <HA>
A:Cross-references: EMBL:X65604
R:Halford, N.G.
submitted to the EMBL Data Library, April 1992
A:Reference number: S24578
A:Accession: S24579
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-61, 'A', 63-513 <HA>
A:Cross-references: EMBL:X65604; NID:G18933; PIN:CAA6554.1; PID:G18934
C:Function:
A:Description: catalyzes the formation of peptidyl-L-serine-phosphate or peptidyl-L-threonine
C:Superfamily: AMP-activated protein kinase; protein kinase homology
C:Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kinase
F.15-272/Domain: protein kinase homology <KIN>
F.23-31/Region: protein kinase ATP-binding motif
F.46,65,143,144/Active site: Lys, Glu, Asp, Lys #status predicted
F.148,157/Binding site: magnesium (Asn, Asp) #status predicted

Query Match	14.4%	Score 273;	DB 1;	Length 513;
-------------	-------	------------	-------	-------------

Best Local Similarity 31.7%; Pred. No. 2e-11;
Matches 79; Conservative 42; Mismatches 102; Indels 26; Gaps 7.

QY 117 HKHVARPEVYLAGQQLAAFFTR--DHGDHSLVRRNRHREDEAAVLFQOMATYALAHQ 175
D 77 HPHLIRVEVETPKQDIFVMEVMEYCNNGELDIYIENGRIQOEBAKRIIFQIILAGVGYCR 136
QY 176 HGLVLRDYLKGRFVPADEBKRLV---LENEDSCVLGPDDSLMDKACPA--YVGPEI 230
D 137 IMVYHRDLKPEVILIDSRVYNVLADFGLSNV-----MRGHEFLKTCGSLNYPAPEI 188
QY 231 LSSRAVSYGKKAADVMSLGVALFTMLAGHYFPQDSEPVLLFGKIRRAVYLLPGLSAPAC 290
D 189 ISSKK--YGFPEVDVMSCGVVLPAALCGSVFDDDNIPSLFRKIKGGTIIYLLPILLSAKD 247
QY 291 LVRCILREPEARLTATGILHAPLRQD--PMPLAPTRSHIMEAAQVVPDGLGDEAREEE 349
D 248 LIPKLTINIDPMKRITFHEIRVHPVFNHILPYLAVPYPKQEOAKMI-----DED 297
QY 350 GDREYVLYYG 358
D 298 ILREVVNLG 306

Search completed: August 24, 2004, 18:53:44
Job time : 21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 18:49:56 ; Search time 12 Seconds
(without alignments)
1553.426 Million cell updates/sec

Title: US-10-070-337-5

Perfect score: 1892
Sequence: 1 MRATPLAAPAGSLSRKKRLK.....GLGDFAAREEGREVVLYG 358

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1888	99.8	358	1 NIPK_HUMAN	Q96YU7 homo sapien
2	1338	70.7	354	1 NIPK_MOUSE	Q8K4K2 mus musculu
3	1324.5	70.3	349	1 NIPK_RAT	Q9WQ66 rattus norv
4	308	16.3	619	1 SNF1_CANTR	Q94168 candida tiro
5	305.5	16.1	646	1 CNK_HUMAN	Q9H4B4 homo sapien
6	303	16.0	620	1 SNF1_CANAL	P52497 candida alb
7	299.5	15.8	631	1 CNK_MOUSE	Q60906 mus musculu
8	294	15.5	794	1 K111_HUMAN	Q8TCD3 homo sapien
9	292	15.4	615	1 CNK_RAT	Q91011 rattus norv
10	284	15.0	1518	1 KKK1_YEAST	Q91011 rattus norv
11	283.5	15.0	611	1 SNF1_CANGA	P34244 saccharomyc
12	280	14.8	535	1 SNF1_YEAST	Q00372 candida gla
13	273.5	14.5	633	1 SNF1_YEAST	Q08997 arabidopsis
14	270.5	14.3	593	1 CDRI_SCHPO	P07334 schizosacch
15	270	14.3	576	1 SNF1_SCHPO	O74536 schizosacch
16	268.5	14.2	533	1 KCCD_RAT	P15791 rattus norv
17	268.5	14.2	736	1 ST29_HUMAN	O81W63 homo sapien
18	266	14.1	542	1 KCCB_MOUSE	P28652 mus musculu
19	264	14.0	622	1 YNA3_CABEL	P45894 caenorhabdi
20	263	13.9	542	1 KCCB_RAT	P08413 rattus norv
21	263	13.9	542	1 KCCB_HUMAN	Q13554 homo sapien
22	262	13.8	1142	1 GIN4_YEAST	Q12263 saccharomyc
23	261.5	13.8	499	1 KCCD_HUMAN	Q13557 homo sapien
24	260.5	13.8	552	1 AAK2_HUMAN	P46466 homo sapien
25	259.5	13.7	552	1 AAK2_MOUSE	O09137 rattus norv
26	259.5	13.7	891	1 KINI_SCHPO	P22967 echinosacch
27	257.5	13.6	786	1 SNIL_HUMAN	P57059 homo sapien
28	257	13.6	502	1 KIL1_SECCB	Q02703 secale cere
29	255	13.5	470	1 CDP1_ARATH	Q06850 arabidopsis
30	251.5	13.3	478	1 KCCA_HUMAN	O9UGM7 homo sapien
31	251.5	13.2	433	1 KCCA_RAT	P11275 rattus norv
32	249.5	13.1	685	1 SNK_HUMAN	Q15831 homo sapien
33	248.5	13.1	685	1 SNK_MOUSE	O9NYJ3 homo sapien

34	246.5	13.0	353	1 ASK2_ARATH	P43292 arabidopsis
35	246	13.0	332	1 AATP_WHEAT	Q02066 triticum ae
36	245	12.9	735	1 K6A1_RAT	O63531 rattus norv
37	243.5	12.9	472	1 KCCG_HUMAN	O13555 homo sapien
38	243.5	12.9	527	1 KCCG_RAT	P11730 rattus norv
39	243.5	12.9	529	1 KCCG_MOUSE	Q92369 mus musculu
40	241.5	12.8	724	1 K6A1_MOUSE	P18653 mus musculu
41	241	12.7	682	1 SNK_MOUSE	P53351 mus musculu
42	241	12.7	1037	1 KCCA_YEAST	P25389 saccharomyc
43	240.5	12.7	478	1 KCCA_MOUSE	P11798 mus musculu
44	240.5	12.7	548	1 AAK1_RAT	P54645 rattus norv
45	240	12.7	542	1 CDP3_ORYSA	P53684 oryza sativ

ALIGNMENTS

RESULT 1
ID NIPK_HUMAN STANDARD. PRT. 358 AA.
AC Q96YU7; Q8TA16; Q9H5M8; Q9HND2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neuronal cell death inducible putative kinase (SKIP3).
GN NIPK OR C20ORP97.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kiss-Toth E., Wyllie D.H., Qvarnstrom E.F., Dower S.K.;
RT Identification of pro-inflammatory cytokine signalling network
RT components by transcription expression screening."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Oca T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT "NEO human cDNA sequencing project."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Bailely O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Garder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Cordy N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Gratlam D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.U., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Lehesvialho M.H., Leverha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McKay K., McMuray A.A.,
RA Milne S.A., Mierry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver A.K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Pratchalingam S.R., Plumb R.W., Ramay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showmreen R., Sims S.,
RA Stuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Symmore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whitlaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilmshurst L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RL "The DNA sequence and comparative analysis of human chromosome 20.";
Nature 414:665-671(2001).

[4]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Cervix, and Muscle;
 RC MEDLINE=22388257; PubMed=12477932;
 RA Strauberg R.L., Feinold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Donald A.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.U., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko I., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smallue D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences".
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: May play an important role in a common pathway leading
 CC to programmed neuronal cell death. Does not appear to function in
 CC the programmed death of non-neuronal cells. May serve as an
 CC endogenous antagonist competing for substrate with functional
 CC kinases that act to promote neuronal cell survival (By
 CC similarity).
 CC -!- SIMILARITY: Contains 1 protein kinase domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF250311; AAK58175.1; -
 CC EMBL; AK026945; BAB15597.1; -
 CC EMBL; AL034548; CAB81634.1; -
 CC EMBL; BC019363; AAB19363.1; -
 CC EMBL; BC027484; AAB27484.1; -
 CC Genew; HGNC:16228; C20orf97.
 CC MIM; 607898; -
 CC InterPro; IPR000719; Prot_kinase.
 CC InterPro; IPR002290; Set_Thr_kinase.
 CC Pfam; PF00069; pkinase; 2.
 CC ProDom; PD000001; Prot_kinase; 1.
 CC SMART; SM00220; S_TKc; 1.
 CC DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
 CC DR Apoptosis.
 CC KW DOMAIN 68 316 PROTEIN KINASE.
 CC FT CONFLICT 84 84 Q -> R (IN REF. 4; AAB27484).
 CC FT CONFLICT 105 105 L -> P (IN REF. 1).
 CC FT CONFLICT 114 114 L -> V (IN REF. 2).
 CC FT CONFLICT 194 195 RR -> DREK (IN REF. 1).
 CC SEQUENCE 358 AA; 39577 MW; CEISF89A81BD63 CRC64;
 SQ
 Query Match 99.8%; Score 1888; DB 1; Length 358;
 Best Local Similarity 99.7%; Pred. No. 7.5e-139;
 Matches 357; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRATPLAPAGSLSKRLLEDDNDTERPVQKARSGPQPLPLPCLPLSPPTAPDRA 60
 DB 1 MRATPLAPAGSLSKRLLEDDNDTERPVQKARSGPQPLPLPCLPLSPPTAPDRA 60
 QY 61 AVATASRLGPPVLLPEPEGRAYOALHCPGTGEYCKYYPQOALAVLEPARLPKHV 120
 DB 61 AVATASRLGPPVLLPEPEGRAYOALHCPGTGEYCKYYPQOALAVLEPARLPKHV 120
 RX 61 AVATASRLGPPVLLPEPEGRAYOALHCPGTGEYCKYYPQOALAVLEPARLPKHV 120

QY 121 ARPTEVLAGTOLLYAFTRTHGDMHSIVTRHRIPEEPAVLFROMATALAHCHQGLVL 180
 DB 121 ARPTEVLAGTOLLYAFTRTHGDMHSIVTRHRIPEEPAVLFROMATALAHCHQGLVL 180
 QY 181 RDLKLCFVFAADRRKKLVLENEDSCVLGPPDLSMDKACPAVYGPETISSASISGK 240
 DB 181 RDLKLCFVFAADRRKKLVLENEDSCVLGPPDLSMDKACPAVYGPETISSASISGK 240
 QY 241 AADVMSIGVALFTMLAGHYFPODESEVPLTFGKIRGAYALPAGLSAPARCLVRLRREP 300
 DB 241 AADVMSIGVALFTMLAGHYFPODESEVPLTFGKIRGAYALPAGLSAPARCLVRLRREP 300
 QY 301 AEBULTAGLILHPLWRDDEPLAPTRSHLWEAAQVDPDGLDEAREEGDREVLVYG 358
 DB 301 AEBULTAGLILHPLWRDDEPLAPTRSHLWEAAQVDPDGLDEAREEGDREVLVYG 358
 RESULT 2
 ID NIPK_MOUSE STANDARD; PRT; 354 AA.
 AC Q8K4K2; Q921E7;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Neuronal cell death inducible putative kinase (TRB-3).
 GN NIPK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22354683; PubMed=12468651;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaido I., Otsu N., Saito R., Suzuki H., Yamana K., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Balarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chochia C., Corbani L.E., Cousins S.,
 RA Dalia E., Dragani T.A., Fletcher C.F., Forrest A., Fraser K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Petrea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.O., Reid J., Ring B.Z., Ringwald M.,
 RA Sanderlin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilting L.G., Wyshaw-Boris A., Yamagisawa M., Yang J., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Kono H., Nakamura S., Sakazune N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arikawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imocani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs".
 RL Nature 420:563-573 (2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rata S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Halys S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodigues J.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield V.S.N., Krzywinski M.T., Skalska U., Smallus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -i- FUNCTION: May play an important role in a common pathway leading
 CC to programmed neuronal cell death. Does not appear to function in
 CC the programmed death of non-neuronal cells. May serve as an
 CC endogenous antagonist competing for substrate with functional
 CC kinases that act to promote neuronal cell survival (By
 CC similarity).
 CC -i- SIMILARITY: Contains 1 protein kinase domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF358668; AAM45476.1; -;
 CC EMBL; AK089931; BAC41002.1; -;
 CC EMBL; BC012955; AAH12955.1; -;
 CC InterPro: IPR000719; Prot kinase.
 CC InterPro: IPR002290; Ser_Thr_kinase.
 CC Pfam; PF00069; pkinase; 2.
 CC ProDom; PD000001; Prot kinase; 1.
 CC SMART; SM00220; S_TKc; 1.
 CC PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
 CC KIM Apoptosis.
 CC FT DOMAIN 68 315 PROTEIN KINASE.
 CC FT CONFLICT 157 157 S -> P (IN REF. 3).
 CC FT CONFLICT 219 219 K -> T (IN REF. 1).
 CC FT CONFLICT 239 264 MISSING (IN REF. 3).
 CC FT CONFLICT 301 354 SERVALGILHPMLREDHGKRVSPQSDREMDQVVDGPGQ
 CC LEDEAGEVGLYG -> CRATGCPGNPLMSIVRGRSPSLS
 CC STY (IN REF. 2).
 CC SEQUENCE 354 AA; 39022 MW; 2CB283FC19F859F CRC64;
 SQ
 Query Match 70.7%; Score 1338; DB 1; Length 354;
 Best Local Similarity 73.5%; Pred. No. 2.5e-96;
 Matches 263; Conservative 31; Mismatches 60; Indels 4; Gaps 2;
 QY 1 MRATPLAAPAGSISRKKRLDLDNLTDPVQRRASGPOPLPCLPLSPPTADPDRAT 60
 DB 1 MRATPLAAPAGSISRKKRLDLDNLTDPVQRRASGPOPLPCLPLSPPTADPDRAT 60
 QY 61 AVATASRLGTYVLEEEGGRAYOALHCPGTGYTCVKYVVOALVLEBYAALPDKHY 120
 DB 61 AVATASRLGTYVLEEEGGRAYOALHCPGTGYTCVKYVVOALVLEBYAALPDKHY 120
 QY 121 ARPTVLAGTQLLYAFRTTHGDMHSLVTRRHIPPEFAAVLFROMATLAHQHGLVL 180
 DB 121 ARPTVLAGTQLLYAFRTTHGDMHSLVTRRHIPPEFAAVLFROMATLAHQHGLVL 180
 QY 181 RDLKLRFPVADRRKKLVLENLEDSCVLTGPPDSDLMDKACPAVYVPELSSRSYSGK 240
 DB 181 RDLKLRFPVADRRKKLVLENLEDSCVLTGPPDSDLMDKACPAVYVPELSSRSYSGK 240

DB 181 RDLKLRFPVADRRKKLVLENLEDSCVLTGPPDSDLMDKACPAVYVPELSSRSYSGK 240
 QY 241 AADVWSIGVALFTWLACHYPPQDSEVYLFGKTRRGYALPAGISAPARCLVRLKREP 300
 DB 241 AADVWSIGVALFTWLACHYPPQDSEVYLFGKTRRGYALPAGISAPARCLVRLKREP 300
 QY 301 AERTATGILHPMLRDPMDLAPTRSHLWMAAVVDDGGLDARREEGDREVLVG 358
 DB 301 SERVALGILHPMLREDHGKRVSPQSDREMDQVVDGPGQLEER--EEG--EVGLYG 354
 RESULT 3
 NIPK_RAT STANDARD; PRT; 349 AA.
 ID NIPK_RAT
 AC NIPK_RAT
 AD 99706;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Neuronal cell death inducible putative kinase.
 GN NIPK.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Neuron;
 RX MEDLINE=99262087; PubMed=10329375;
 RA Mayumi-Macuda K., Kojima S., Suzuki H., Sakata T.;
 RT "Identification of a novel kinase-like gene induced during neuronal
 RT cell death.";
 RL Biochem. Biophys. Res. Commun. 258:260-264(1999).
 CC -i- FUNCTION: May play an important role in a common pathway leading
 CC to programmed neuronal cell death. Does not appear to function
 CC in the programmed death of non-neuronal cells. May serve as an
 CC endogenous antagonist competing for substrate with functional
 CC kinases that act to promote neuronal cell survival.
 CC -i- TISSUE SPECIFICITY: Detected only in the lung. Not detected in the
 CC heart, brain, spleen, liver, skeletal muscle, kidney and testis.
 CC -i- INDUCTION: Expression induced during programmed cell death evoked
 CC in neuronal cells by NGF-depletion.
 CC -i- SIMILARITY: Contains 1 protein kinase domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB020967; BAA77582.1; -;
 CC InterPro: IPR000719; Prot kinase.
 CC Pfam; PF00069; pkinase; 2.
 CC ProDom; PD000001; Prot kinase; 1.
 CC PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
 CC KIM Apoptosis.
 CC FT DOMAIN 63 310 PROTEIN KINASE.
 CC FT CONFLICT 301 354 SERVALGILHPMLREDHGKRVSPQSDREMDQVVDGPGQ
 CC LEDEAGEVGLYG -> CRATGCPGNPLMSIVRGRSPSLS
 CC STY (IN REF. 2).
 CC SEQUENCE 349 AA; 38602 MW; 3050F9BF9E346D815 CRC64;
 SQ
 Query Match 70.0%; Score 1324.5; DB 1; Length 349;
 Best Local Similarity 72.9%; Pred. No. 2.7e-95;
 Matches 261; Conservative 28; Mismatches 60; Indels 9; Gaps 3;
 QY 1 MRATPLAAPAGSISRKKRLDLDNLTDPVQRRASGPOPLPCLPLSPPTADPDRAT 60
 DB 1 MRATPLAAPAGSISRKKRLDLDNLTDPVQRRASGPOPLPCLPLSPPTADPDRAT 60
 QY 61 AVATASRLGTYVLEEEGGRAYOALHCPGTGYTCVKYVVOALVLEBYAALPDKHY 120
 DB 61 AVATASRLGTYVLEEEGGRAYOALHCPGTGYTCVKYVVOALVLEBYAALPDKHY 120
 QY 121 ARPTVLAGTQLLYAFRTTHGDMHSLVTRRHIPPEFAAVLFROMATLAHQHGLVL 180
 DB 121 ARPTVLAGTQLLYAFRTTHGDMHSLVTRRHIPPEFAAVLFROMATLAHQHGLVL 180

```

Db      116 ARPEVLIGSGLTYLFTFKTHGDLHSIVSRSGIPEPEAAALFRQMSAVAHCHKHGLIL 175
Qy      181 RDLKLCFVFADRRKRLVLENLEDSCLVNGPDSDIMDKACPAVYGPETISSASYSCK 240
Db      176 RDLKLRFPVSNCRRTKLVLENLEDSCLVNGPDSDIMDKACPAVYGPETISSASYSCK 235
Qy      241 AADVMSIGVALFTMLAGHYPPQDSEPVLLFGKIRGAYALPAGLSAPARCIVRCILRRREP 300
Db      236 AADVMSIGVALFTMLAGHYPPQDSEPVLLFGKIRGAYALPAGLSAPARCIVRCILRRREP 295
Qy      301 AERLTATGIIILHPLRQDPMPLAFTRSHLMEAAQVVDGILGDAREEGDREVIYVG 358
Db      296 SERLVALGIIILHPLRQDPMPLAFTRSHLMEAAQVVDGILGDAREEGDREVIYVG 349

```

RESULT 4

```

SNP1_CANTR STANDARD; PRT; 619 AA.
ID 094168;
AC 094168;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Carbon catabolite derepressing protein kinase (EC 2.7.1.-).
GN SNP1.
OS Candida tropicalis (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5482;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RA Kanal T., Ogawa K., Ueda M., Tanaka A.;
RT "Genetic evaluation of the function of SNP1 in Candida tropicalis.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Essential for release from glucose repression. It
CC interacts and has functional relationships to the regulatory
CC protein SNF4. Could phosphorylates CAr8 (By similarity).
CC -1- SUBCELLULAR LOCATION: Associated with the nuclear membrane (By
CC similarity).
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. SNP1
CC subfamily.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL; AB024535; BAA75889.1; -.
DR HSSP; O63450; 1A06.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_Thr_kin_AS.
DR InterPro; IPR002290; Ser_Thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR Transferrase; Serine/threonine-protein kinase; ATP-binding;
DR Phosphorylation; Carbohydrate metabolism; Nuclear protein.
KW DOMAIN
FT DOMAIN 12 29
FT NP_BIND 58 66
FT BINDING 81 81
FT ACT_SITE 174 174
FT ACT_SITE 207 207
FT MOD_RES 619 AA; 70323 MW; 0FCF1FC3DCE706D7 CRC64;
SQ SEQUENCE

```

```

Query Match 16.3%; Score 308; DB 1; Length 619;
Best Local Similarity 30.9%; Pred. No. 2.1e-16;

```

```

Matches 95; Conservative 48; Mismatches 128; Indels 36; Gaps 12;
Qy      40 QRLPPECLPLSPPTADPRATAVATASRLGPEYLLER-EEG--GRAYQALHCPGTETC 96
Db      32 QPAPQ--IPIDPNVN-----ANRIGRYQIKTLGEGSFQKVALAGHVGQGVALL 80
Qy      97 KYVEQVQALVLE-----PYARLPKHVARPFEVLAAGTOLLYAFRTKGDHSL 147
Db      81 KIIN-RKTLKSDMQGVVEREISTYRLRHPHIKIVDKVSKDEILMTVIEFPAKEDFDY 139
Qy      148 VRTNRRIPEPEAAVLFROMATALACHQHGILVLDLCLCFVFADRRKRLVLENLEDSCL 207
Db      140 IVRGKKMPDEARFPQGIILAAVEYCHRKIKIVRDLKPEVLLDDQLVNKTADPGL--SN 197
Qy      208 VLTGPDSDIMDKAC--PAYVGPETISSRASYSCKADVMSIGVALFTMLAGHYPPQDSE 265
Db      198 IMT---DGNFLKTSQSGSPNYAAPEVISGKL-YAGPEVDVMSGVIIYVLMCGRLPFDEF 253
Qy      266 PVLFGKIRGAYALPAGLSAPARCIVRCILRRREPARELTAATGIIILHPLRQDPM---L 322
Db      254 IPLAFKXISGVYTLDPNVISPGAKHLLTRMLVVPNPLRITHEIMEDEWFKOD-MPDYIL 312
Qy      323 APTRSHL 329
Db      313 PPDLSKI 319

```

RESULT 5

```

CNK_HUMAN STANDARD; PRT; 646 AA.
ID 09H434; O15767;
AC 09H434; O15767;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Cytochrome-inducible serine/threonine-protein kinase (EC 2.7.1.37) (FGF-
DE inducible kinase) (Proliferation-related kinase).
GN CNK OR FNK OR PRK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RA TISSUE=Embryo.
RC MEDLINE=20493044; PubMed=11039900;
RA Holtrich U., Wolf G., Yuan J., Bereiter-Hahn J., Karn T., Weiler M.,
RA Kaesmann G., Rehli M., Andreesen R., Kaufmann M., Kuhl D.,
RA Strebhardt K.;
RT "Adhesion induced expression of the serine/threonine kinase Fnk in
RT human macrophages.";
RT Oncogene 19:4832-4839(2000).
RN [2]
RP SEQUENCE OF 28-646 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=96325053; PubMed=8702627;
RA Li B., Ouyang B., Pan H., Reissmann P.T., Slamon D.J., Arceci R.,
RA Lu L., Dai W.;
RT "Prk, a cytokine-inducible human protein serine/threonine kinase whose
RT expression appears to be down-regulated in lung carcinomas.";
RT J. Biol. Chem. 271:19402-19408(1996).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=98019242; PubMed=9353331;
RA Ouyang B., Pan H., Lu L., Li U., Stambrook P., Li B., Dai W.;
RT "Human Prk is a conserved protein serine/threonine kinase involved in
RT regulating M phase functions.";
RN J. Biol. Chem. 272:28646-28651(1997).
RN [4]
RX CHARACTERIZATION.
RX MEDLINE=20027391; PubMed=10557092;
RA Ouyang B., Li W., Pan H., Meadows J., Hoffmann I., Dai W.;
RT "The physical association and phosphorylation of Cdc25C protein
RT phosphatase by Prk.";

```

RL Oncogene 18:6029-6036(1999).
 CC -1- FUNCTION: Serine/threonine protein kinase involved in regulating M
 CC phase functions during the cell cycle. May also be part of the
 CC signaling network controlling cellular adhesion. In vitro, is able
 CC to phosphorylate CDC25C and casein.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -1- SUBUNIT: Binds to the calcium/integrin-binding protein (CIB). This
 CC interaction probably occurs via the POLO-box domain.
 CC -1- SUBCELLULAR LOCATION: Membrane-associated.
 CC -1- TISSUE SPECIFICITY: Transcripts are highly detected in placenta,
 CC lung, followed by skeletal muscle, heart, pancreas, ovaries and
 CC kidney and weakly detected in liver and brain. May have a short
 CC half-life, in cells of hematopoietic origin, strongly and
 CC exclusively detected in terminally differentiated macrophages.
 CC Transcript expression appears to be down-regulated in primary
 CC lung tumor.
 CC -1- INDUCTION: Cytokine and cellular adhesion trigger FAK induction.
 CC -1- PTM: PHOSPHORYLATED AS CELLS ENTER MITOSIS AND DEPHOSPHORYLATED AS
 CC CELLS EXIT MITOSIS (BY SIMILARITY).
 CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
 CC CDC5/Polo subfamily.
 CC -1- SIMILARITY: Contains 2 POLO box domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AJ293866; CAC10659.1; -.
 CC DR EMBL: U56998; AAC50637.1; ALU_INIT.
 CC DR Genew; HGNC:2154; CNK.
 CC DR GK; Q94B4; -.
 CC DR MIM; 602913; -.
 CC DR GO; GO:0004674; F:protein serine/threonine kinase activity; TAS.
 CC DR GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
 CC DR GO; GO:0000074; P:regulation of cell cycle; TAS.
 CC DR InterPro; IPR000959; POLO box.
 CC DR InterPro; IPR000719; Prot_kinase.
 CC DR InterPro; IPR008271; Ser_thr_kin_AS.
 CC DR InterPro; IPR002290; Ser_thr_pkinase.
 CC DR Pfam; PF00069; pkinase; 1.
 CC DR Pfam; PF00659; POLO box; 2.
 CC DR ProDom; PD000001; Prot_kinase; 1.
 CC DR SMART; SM00220; S_TKc; 1.
 CC DR PROSITE; PS50078; POLO_BOX; 2.
 CC DR PROSITE; PS50107; PROTEIN_KINASE_ATP; 1.
 CC DR PROSITE; PS5011; PROTEIN_KINASE_DOM; 1.
 CC DR PROSITE; PS00108; PROTEIN_KINASE_SR; 1.
 CC KW Transferase; Serine/threonine-protein kinase; ATP-binding; Repeat;
 CC phosphorylation.
 CC FT DOMAIN 62 314 PROTEIN KINASE.
 CC FT NP BIND 68 76 ATP (BY SIMILARITY).
 CC FT BINDING 91 91 ATP (BY SIMILARITY).
 CC FT ACT SITE 185 185 BY SIMILARITY.
 CC FT DOMAIN 470 537 POLO BOX 1.
 CC FT DOMAIN 367 637 POLO BOX 2.
 CC FT CONFLICT 99 99 V -> A (IN REF. 2).
 CC FT CONFLICT 353 353 V -> G (IN REF. 2).
 CC FT CONFLICT 419 419 H -> D (IN REF. 2).
 CC FT CONFLICT 464 470 FSEWGVF -> VSKWVDY (IN REF. 2).
 CC FT CONFLICT 522 522 R -> P (IN REF. 2).
 CC FT SEQUENCE 646 AA; 71789 MW; C20147CD0F8AB4 CRC64;
 QY Query Match 16.1%; Score 305.5; DB 1; Length 646;
 QY Best Local Similarity 28.9%; Pred. No. 3.5e-16;
 QY Matches 101; Conservative 46; Mismatches 134; Indels 67; Gaps 15;
 DB 11 RPFQRTAAATAPAGPBP--PFSALRGPELEMLAGLTPSDPGRLLTPRSGRTYLVKGRLL 68

QY 69 GPVLLLEPERGG--RAYQALHCPRTGYTCKVYVQDALVLEPYARLP----- 115
 DB 69 G-----KGGFARCEADTDETSAYAVKVPQSR---VVKHOREKINIELHRDL 117
 QY 116 PHKVARPTVLAAGTOLLYAFETR-THGDMHSLVTRHRIPEPAAYLFROMATALACH 174
 DB 118 QHRITVPSHHFFEDADNIVYFLFELCSKSLAHMKAHHTLLEPRVRYKRIISGLKYLH 177
 QY 175 QHGLVLDLKLCPFPVADRERKKVLLENEDSCVLTPGDDSLMDKAC--PAYGPEILS 232
 DB 178 QRLTHHDLKLGNFIT--ENMELKVGDFGLAARLEPEQR--KKTICGTNNYVAPEVLL 233
 QY 233 SRASYSGKADVWSLGVALFTMLAGHPQDSEVLLFGKRRGAYALPAGISAPRGLV 292
 DB 234 RQG--HGPEADVWSLGCVMYTLTGSPPEFADAKETRYCRICKQVHYTLPAISLPRQDL 291
 QY 293 RCLRRPFAERLNTANGTLHPMLRQ---DPMPIAPTRSHLMEAAQVVPD 338
 DB 292 AAILRASPDRPSIDQLHRHDFFTKGYTPDKPLIS-----SCVTVPD 333

RESULT 6

SNF1_CANAL STANDARD; PRT; 620 AA.
 ID SNF1_CANAL
 AC P52497; Q00309;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Carbon catabolite derepressing protein kinase (EC 2.7.1.-).
 GN SNF1.
 OS Candida albicans (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=5476;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 3254 / B-311;
 RX MEDLINE=98053924; PubMed=9393775;
 RA Pelter R., Kwon-Chung K.U.;
 RT "A gene homologous to *Saccharomyces cerevisiae* SNF1 appears to be
 RT essential for the viability of *Candida albicans*,"
 RL Infect. Immun. 65:4909-4917(1997).
 RN [2]
 RP SEQUENCE OF 7-620 FROM N.A.
 RC STRAIN=ATCC 3254 / B-311;
 RA Pelter R., Kwon-Chung K.U.;
 RL Submitted (MAR-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: Essential for release from glucose repression. It
 CC interacts with and has functional relationship to the regulatory
 CC protein SNF4. Could phosphorylates Cat8 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Associated with the nuclear membrane (By
 CC similarity).
 CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. SNF1
 CC subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L78129; AAB46643.1; -.
 CC DR EMBL: L39263; AAB92456.1; -.
 CC DR HSRP; Q63450; 1A06.
 CC DR InterPro; IPR000719; Prot_kinase.
 CC DR InterPro; IPR008271; Ser_thr_kin_AS.
 CC DR InterPro; IPR002290; Ser_thr_pkinase.
 CC DR Pfam; PF00069; pkinase; 1.
 CC DR ProDom; PD000001; Prot_kinase; 1.
 CC DR SMART; SM00220; S_TKc; 1.

DR PROSITE, PS00107; PROTEIN KINASE ATP, 1.
 DR PROSITE, PS00108; PROTEIN KINASE ST, 1.
 DR PROSITE, PS50011; PROTEIN KINASE DOM, 1.
 KM Transferrase; Serine/threonine-protein kinase; ATP-binding;
 KM Phosphorylation; Carbohydrate metabolism; Nuclear protein.
 FT DOMAIN 16 29
 FT NP BIND 53 305
 FT BINDING 59 67
 FT ACT SITE 82 82
 FT MOD RES 175 175
 FT MOD RES 208 208
 FT CONFLICT 208 208
 FT CONFLICT 228 234
 FT CONFLICT 242 255
 FT CONFLICT 387 387
 FT CONFLICT 416 416
 FT CONFLICT 494 494
 SQ SEQUENCE 620 AA; 70005 MW; 1806c52b5061d2b CRC64;
 Query Match 16.0%; Score 303; DB 1; Length 620;
 Best Local Similarity 29.4%; Pred. No. 5,2e-16;
 Matches 95; Conservative 55; Mismatches 131; Indels 42; Gaps 12;
 QY 48 LPLSPPLAPRPAVAATASRLGYPVLLP-EEG--GRAYALHCPGTGYTCRVYPOEA 104
 DB 38 VPIDPAMP-----ANRIGRYQLTKLGSGFGVKLAQHLGQKVALKIIN-RTK 88
 QY 105 LAYLE-----PYARLPHKHVAPRTFVLAAGTQLVAFRTIGDMSLVTRTHRP 155
 DB 89 LMSDMQGRVERISTURLRHRPHITKLVDVSKDEIIVIEFAKELEFDYVQSGKMP 148
 QY 156 EPEAAVLFRMATATALACHQHLVLRDLKCRFVFADREKKIVLENLEDSCVLTGPDOS 215
 DB 149 EDEARFEOQIIAAYEYCHRHKIVHMDIKRENLLDDQNVAKIADGL--SNMT---DG 203
 QY 216 LMDKHC--PAYV-GEIILSRASISGKADWLSGLVFTMLAGHYPFQDSEPVLLFGK 272
 DB 204 NPLKTSIGSPNVPAPAEVVISGKL-VAGPEVDVMSAGVILYVMLCGRLPFDEPIPALFKK 262
 QY 273 IRRGAVALPAGSAPARCVRCLAREPARELTAATGILLAPMKRQPMPLAPRSHLMA 332
 DB 263 ISNGVTLTNRYSAGAKHLITRMLVYNPLNRIITHEIMEDWFKOD-MP----- 310
 QY 333 AQVVPDGLGLDEAREEGDREVV 355
 DB 311 DYLPLPDLSKNKNSKIDVEDVI 333
 RESULT 7
 CNK_MOUSE STANDARD; PRT; 631 AA.
 ID CNK_MOUSE Q60806; Q60822; Q9R009;
 AC Q60806; Q60822; Q9R009;
 DT 16-OCT-2001 (Rel. 40. Created)
 DT 16-OCT-2001 (Rel. 40. Last sequence update)
 DT 28-FEB-2003 (Rel. 41. Last annotation update)
 DE Cytokine-inducible serine/threonine-protein kinase (EC 2.7.1.37) (RGE-
 inducible kinase).
 GN CNK OR FNK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OK NCBI_taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX STRAIN=NIH Swiss;
 RC MEDLINE=95247749; PubMed=7730342;
 RA Donohue P.V.; Alberts G.F.; Guo Y.; Winkles J.A.;
 RT "Identification by targeted differential display of an immediate early
 RT gene encoding a putative serine/threonine kinase."
 RL J. Biol. Chem. 270:10351-10357(1995).
 RN [2]
 RP SEQUENCE OF 333-437 FROM N.A. (ISOFORM 2).

RC STRAIN=NIH Swiss;
 RA Kauselmann G.; Weiler M.; Kuhl D.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP FUNCTION, AND PHOSPHORYLATION.
 RX MEDLINE=98343954; PubMed=9677325;
 RA Chase D.; Peng Y.; Haneshe B.; Winkles J.A.; Longo D.L.; Ferris D.K.;
 RT "Expression and phosphorylation of fibroblast-growth-factor-inducible
 RT kinase (Ftk) during cell-cycle progression."
 RL Biochem. J. 333:655-660(1998).
 CC -1- FUNCTION: Serine/threonine protein kinase involved in regulating M
 CC phase functions during the cell cycle. May also be part of the
 CC signaling network controlling cellular adhesion. In vitro, is able
 CC to phosphorylate CDC35C and casein (By similarity).
 CC -1- SUBUNIT: Binds to the calcium/integrin-binding protein (CIB). This
 CC interaction probably occurs via the POLO-box domain (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Membrane-associated (By similarity).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q60806-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q60806-2; Sequence=VSP_004927;
 CC -1- TISSUE SPECIFICITY: Expressed in skin.
 CC -1- PTM: PHOSPHORYLATED AS CELLS ENTER MITOSIS AND DEPHOSPHORYLATED AS
 CC CELLS EXIT MITOSIS.
 CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
 CC CDC5/Polo subfamily.
 CC -1- SIMILARITY: Contains 2 POLO box domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U21392; AAC52191.1; -;
 CC EMBL; U22434; AAC52192.1; -;
 CC EMBL; AF136586; AAF08369.1; -;
 CC PIR; A57286; A57286.
 CC MGI; MGI:109604; Cnk.
 CC InterPro; IPR000959; POLO_box.
 CC InterPro; IPR000719; Prot_kinase.
 CC InterPro; IPR008271; Ser_thr_kin AS.
 CC InterPro; IPR002290; Ser_thr_kinase.
 CC Pfam; PF00659; pkinase; 1.
 CC Pfam; PF00659; POLO_box; 2.
 CC ProDom; PD000001; Prot_kinase; 1.
 CC SMART; SM00220; S_TKc; 1.
 CC PROSITE; PS50078; POLO_BOX; 2.
 CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 CC PROSITE; PS50011; PROTEIN KINASE DOM; 1.
 CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KM Transferrase; Serine/threonine-protein kinase; ATP-binding; Repeat;
 KM Phosphorylation; Alternative splicing.
 FT DOMAIN 63 315
 FT NP BIND 69 77
 FT BINDING 92 92
 FT ACT SITE 186 186
 FT DOMAIN 455 518
 FT DOMAIN 552 622
 FT VASPLLC 373 373
 FT CONFLICT 386 386
 FT SEQUENCE 631 AA; 70012 MW; 20857341870BD1D2 CRC64;
 Query Match 15.8%; Score 299.5; DB 1; Length 631;
 Best Local Similarity 28.3%; Pred. No. 9.9e-16;
 Matches 98; Conservative 48; Mismatches 133; Indels 67; Gaps 14;


```

Db 71 ITGQKVAIKIVNEKLSSEVLMKVEREIALI-----KLIHPHVLLKLDVYENKKYLYVL 126
QY 138 TR--THGDMHSLVTRRHRIPEPEAAVLFROMATALAHCHQGLVLRDCLKRFVADDERK 136
Db 127 EHVSGELPFIYLVKGRHLTKRKARKFRQVSAIDFCHASISICHRDLKPNLL--DEKN 184
QY 197 KLVLENEDSCVLTPGDDSLMDKAC--PAYVGEIILSRASYSKGADVWSLGVAFIM 254
Db 185 NRIADFGMAVLQVG--DSILET--SCGSPHYACPEVTKGE-KYDGRADWMSGCVLLFAL 240
QY 255 LAGHYPRQDSEPVLLFKIRGAYALPAGISAPARCVCVRLKREPARELTATGILLHPW 314
Db 241 LVGALPDDDNRLQRLKRVKRGVFMHFTLPDCCSLRGMIEVEPEKRLSLBQIKHPW 300
QY 315 L--RQDPMPV--APTRSHLMEAAQVVPDGLGD-----EARE 347
Db 301 YLGGKHPPDCLEPAPGR--VAMSLSPNGELDPVLESMAISLGCDFRDLRHLRS 357
QY 348 EEDREVVLY 357
Db 358 EENOEKMTY 367

```

RESULT 9

CKN_RAT STANDARD; PRT; 615 AA.

```

AC Q9R011;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cytochrome-inducible serine/threonine-protein kinase (EC 2.7.1.37) (FGF-
DE inducible kinase) (Fragment).
GN CKN OR FNK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=9452760; PubMed=10523297;
RA Kauselmann G., Weller M., Wulff P., Jessberger S., Konietzko U.,
RA Scatford J., Stambli U., Bereliter-Hahn J., Streibhardt K., Kuhl D.,
RT "The polo-like protein kinases Fnk and Snk associate with a Ca(2+)-and
RT integrin-binding protein and are regulated dynamically with synaptic
RT plasticity."
RT EMBO J. 18:5528-5539(1999).
RL
CC -!- FUNCTION: Serine/threonine protein kinase involved in regulating M
CC phase functions during the cell cycle. May also be part of the
CC signaling network controlling cellular adhesion. In vitro, is able
CC to phosphorylate CDC25C and casein (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- SUBUNIT: Binds to the calcium/integrin-binding protein (CIB). This
CC interaction probably occurs via the POLO-box domain.
CC -!- SUBCELLULAR LOCATION: WHEN INDUCED, IT TRANSLOCATES INTO THE
CC DENDRITES OF ACTIVATED NEURONS.
CC -!- TISSUE SPECIFICITY: Constitutively expressed in post-mitotic
CC neurons.
CC -!- INDUCTION: By the intense activity associated with seizures.
CC -!- PTM: PHOSPHORYLATED AS CELLS ENTER MITOSIS AND DEPHOSPHORYLATED AS
CC CELLS EXIT MITOSIS (By similarity).
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC CDGS/Polo subfamily.
CC -!- SIMILARITY: Contains 2 POLO box domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; AF136584; AAF08367.1; -.

```

```

DR InterPro: IPR000959; POLO box.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR008271; Ser_thr_pkin_AS.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00659; pkinase; 1.
DR Pfam: PF00659; POLO box; 2.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR PROSITE: PSS0078; POLO BOX; 2.
DR PROSITE: PSS0107; PROTEIN KINASE_ATP; 1.
DR PROSITE: PSS0011; PROTEIN KINASE_DOM; 1.
DR PROSITE: PSS0108; PROTEIN KINASE_ST; 1.
DR TRANSFERASE: Serine/threonine-protein kinase; ATP-binding; Repeat;
KM Phosphorylation.
FT NON_TER 1 1
FT DOMAIN 31 283 PROTEIN KINASE.
FT NE_BIND 37 45 ATP (BY SIMILARITY).
FT BINDING 60 60 ATP (BY SIMILARITY).
FT ACT_SITE 154 154 BY SIMILARITY.
FT DOMAIN 440 503 POLO BOX 1.
FT DOMAIN 537 607 POLO BOX 2.
FT NON_TER 615 615
SQ SEQUENCE 615 AA; 68800 MW; 06584C229B5D71BD CRC64;

```

Query Match 15.4%; Score 292; DB 1; Length 615;
 Best Local Similarity 28.5%; Pred. No. 3.6e-15;
 Matches 94; Conservative 47; Mismatches 127; Indels 62; Gaps 13;

```

QY 41 PRLPCLILSPPTADRAATAVATASRLGPVLLPEEG-----GRAYQL 86
Db 3 PSEPEV--LAPQVPD-----ASRL-----ITDRSGRTYIKRLLKGGFARCYEAT 49
QY 87 HCPPTGETYCKYYPVGEALAVLEPYARL-----PKHVAPPEVLAAGTOLLYA 135
Db 50 DRETSAVAVKVIPOGR--VAKPHQREKINEIELHARDLQHRHIVRFSHHEEDADNIIYI 106
QY 136 PFTR--THGDMHSLVTRRHRIPEPEAAVLFROMATALAHCHQGLVLRDCLKRFVADDER 194
Db 107 FLELCRSKSLAHVWAKRHILPEPVRYIRQLSLGKTYHOGSILHRDLKGNFTTD-- 164
QY 195 RKVLLENEDSCVLTPGDDSLMDKAC--PAYVGEIILSRASYSKGADVWSLGVAFIM 252
Db 165 NMELKVGDPGLAARLEPEQR--KTIICGTPVVAPEVLLRQG--HGPEADVWSIGCVWY 220
QY 253 TMLAGHYPRQDSEPVLLFKIRGAYALPAGISAPARCVCVRLKREPARELTATGILLHP 312
Db 221 TLLCGSPPEETADLETKYRKIVHYTLPASLSLPARQLAALTRASPDRPSIEQILRH 280
QY 313 PWLRQ---DPMPLAPTRSHLMEAAQVVPD 338
Db 281 DFTTGYTPDRLPVS-----SCVYVPD 302

```

RESULT 10

KKK1_YEAST STANDARD; PRT; 1518 AA.

```

AC P3424;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Probable serine/threonine-protein kinase YKL101W (EC 2.7.1.-).
GN YKL101W OR YKL453.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxId=4932;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=94078677; PubMed=8256524;
RA Paller C., Valens M., Puzos V., Fukuhara H., Cheret G., Sor F.,
RA Bolotin-Fukuhara M.;
RT "DNA sequence analysis of a 17 kb fragment of yeast chromosome XI

```

```

RT physically localizes the MRL1 gene and reveals eight new open reading
RT frames, including a homologue of the KIN1/KIN2 and SNF1 protein
RL Yeast 9:1149-1155(1993).
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIML
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X71133; CAAS0456.1; -.
DR EMBL; Z28101; CAAB1941.1; -.
DR PIR; S37928; S37928.
DR HSSP; Q63450; 1A06.
DR Germonline; 139857; -.
DR SGD; S0001584; HSL1.
DR GO; GO:0005935; C:bud neck; IDA.
DR GO; GO:0005940; C:septin ring; IDA.
DR GO; GO:0004672; F:protein kinase activity; IDA.
DR GO; GO:0000086; P:G2/M transition of mitotic cell cycle; IGI.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IDA.
DR GO; GO:0000074; P:regulation of cell cycle; IMP.
DR GO; GO:0000135; P:septin checkpoint; IGI.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_kin_AS.
DR InterPro; IPR001245; Tyr_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
KM Hypothetical protein; Transferase; Serine/threonine-protein kinase;
KW ATP-binding.
FT DOMAIN 81 369 PROTEIN KINASE.
FT NP_BIND 87 95 ATP (BY SIMILARITY).
FT BINDING 110 110 ATP (BY SIMILARITY).
FT ACT_SITE 239 239 BY SIMILARITY.
SQ SEQUENCE 1518 AA; 169592 MW; 803F84F7531241DD CRC64;

Query Match 15.0%; Score 284; DB 1; Length 1518;
Best Local Similarity 26.4%; Pred. No. 4.3e-14;
Matches 85; Conservative 51; Mismatches 128; Indels 58; Gaps 9;

```

```

DB 356 KRITTOELIKHPLIKKYDDLPUV 377
RESULT 11
SNF1_CANGA
ID SNF1_CANGA STANDARD; PRT; 611 AA.
AC Q00372;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Carbon catabolite derepressing protein kinase (EC 2.7.1.-).
GN SNF1.
OS Candida glabrata (Yeast) (Torulopsis glabrata).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5478;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCCLS84;
RX MEDLINE=97101049; PubMed=8945576;
RA Pether R., Kwon-Chung K.J.;
RT "Disruption of the SNF1 gene abolishes trehalose utilization in the
RT pathogenic yeast Candida glabrata.";
RL Infect. Immun. 64:5269-5273(1996).
CC -!- FUNCTION: Essential for release from glucose repression. It
CC interacts and has functional relationship to the regulatory
CC protein SNF4. Could phosphorylates Ctr8 (By similarity).
CC -!- SUBCELLULAR LOCATION: Associated with the nuclear membrane (By
CC similarity).
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. SNF1
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L78130; AAB48642.1; -.
DR HSSP; P24941; IHCL.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_kin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRODOM; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
KM Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation; Carbohydrate metabolism; Nuclear protein.
FT DOMAIN 6 17 POLY-HIS.
FT NP_BIND 39 290 PROTEIN KINASE.
FT BINDING 45 53 ATP (BY SIMILARITY).
FT ACT_SITE 68 68 ATP (BY SIMILARITY).
FT MOD_RES 161 161 BY SIMILARITY.
FT MOD_RES 194 194 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 611 AA; 70049 MW; 89E17812A4900CD0 CRC64;

Query Match 15.0%; Score 283.5; DB 1; Length 611;
Best Local Similarity 27.9%; Pred. No. 1.6e-14;
Matches 90; Conservative 54; Mismatches 136; Indels 43; Gaps 10;

```

QY 165 QMATALAHCHQHLVLDLKLCEFFVADREKKVLLENLEDSCVLTGPDSDMKAC-- 222
 Db 144 QIIISAVECHHRKIVHDLKPEMLLDEHNLVXIADGSL--SNMT---DSEFLKTSOGS 198
 QY 223 PAYVGPILSSRASYSGKADVMSLVALFTMLAGHPPODSBPVLLFGKTRGAYALPA 282
 Db 199 PNVAPEVITSGKL-VAGEPVDVMSGCVLTYVLMCRRLPFDEDSIPVLFKNISNGVYTLPK 257
 QY 283 GLSAPARCLVYCLLRPAERLTATGILLHPWLRQDMPPLAPTSIHLMEAAQVYDDGL-- 340
 Db 258 FLSPGASDLIKRMLVLPARKISIHETMODEMFVYD-----LAEYLVPODLKQ 305
 QY 341 -----GLDEAREEGEDREV 355
 Db 306 QEOFNKKSQNEENVEEIDDEWV 328

RESULT 12
 ID K110_ARATH STANDARD; PRT; 535 AA.
 AC Q38997; 004728; Q39076; Q8RWD2;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DE SNRP1-related protein kinase KIN10 (EC 2.7.1.-) (AKIN10).
 GN KIN10 OR SKIN10 OR AT3G01090 OR T4P13.22.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopses.
 NC NCBL_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=93013041; PubMed=1339373;
 RA 1e Gen L., Thomas M., Bianchi M., Halford N.G., Kreis M.;
 RT "Structure and expression of a gene from Arabidopsis thaliana
 RL encoding a protein related to SNRP1 protein kinase."
 RN Gene 120:249-254(1992).
 RL [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX STRAIN=cv. Columbia;
 RA Lessard P., Kreis M., Thomas M.;
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX STRAIN=cv. Columbia;
 RA MEDLINE=21016720; PubMed=1130713;
 RA Salanoubat M., Lemcke K., Rieger W., Ansoerge W., Unseld M.,
 RA Faremann B., Valle G., Bloeker H., Perez-Alonso M., Obermaier B.,
 RA Deleney M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,
 RA De Simone V., Choisme N., Artiguenave F., Robert C., Brotier P.,
 RA Wincker P., Catolico L., Weissbach J., Saurin W., Quetier F.,
 RA Schaefer M., Meiller-Auer S., Gabel C., Fuchs M., Benes V.,
 RA Wurmbach E., Drzonek H., Erfle H., Jordan N., Bangert S.,
 RA Wiedemann R., Krenz H., Voss H., Holland R., Brandt P., Nyakatura G.,
 RA Vezzi A., D'Angelo M., Pallavicini A., Toppi S., Simionati B.,
 RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordstiek G.,
 RA Reichelt J., Schafte M., Schuen O., Barques M., Teol J., Clement J.,
 RA Navarro P., Collado C., Perez-Perez A., Ottenwalder B., Duchemin D.,
 RA Cooke R., Lande M., Berger-Liauz C., Purnelle B., Masny D.,
 RA de Haan M., Maarse A.C., Alataiz J.-P., Cotter B., Casasuberta E.,
 RA Montfort A., Argiliron A., Flores M., Iguori R., Vitale D.,
 RA Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
 RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
 RA Rooney T., Rizzo M., Walts A., Uteback T., Fujii C.Y., Shea T.P.,
 RA Cressy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
 RA Pai G., Miltcher J., Sellers P., Gill J.E., Feldlyum T.V.,
 RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
 RA Fraser C.M., Kaneo T., Nakamura Y., Sato S., Kato T., Asamizu E.,
 RA Sasamoto S., Kimura T., Ideawa K., Kawashina K., Kishida Y.,
 RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakayama S., Nakazaki N., Shino S., Takeuchi C., Wada T.,
 RA Watanabe A., Yamada M., Yasuda M., Tabata S.;

RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 408:820-822(2000).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX STRAIN=cv. Columbia;
 RX MEDLINE=22954850; PubMed=14593172;
 RA Yamada K., Lim U., Dale J.M., Chen H., Shim P., Palm C.J.,
 RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
 RA Katlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
 RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
 RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Atiyana K., Anesi Y.,
 RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carrinci P.,
 RA Chao O., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
 RA Hayashizaki Y., Johnson-Hopson C., Huan V.W., Iida K., Karnes M.,
 RA Khan S., Koeseema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
 RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
 RA Satou M., Tame R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
 RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.,
 RT "Empirical analysis of transcriptional activity in the Arabidopsis
 RT genome.";
 RL Science 302:842-846(2003).
 RN [5]
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORM 2).
 RX STRAIN=cv. Columbia;
 RX MEDLINE=95115691; PubMed=7816049;
 RA 1e Gen L., Thomas M., Kreis M.;
 RT "Gene density and organization in a small region of the Arabidopsis
 RT thaliana genome.";
 RL Mol. Gen. Genet. 245:390-396(1994).
 CC -1- FUNCTION: May play an important role in a signal transduction
 CC cascade regulating gene expression and carbohydrate metabolism in
 CC higher plants.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q38997-1; Sequence=displayed;
 CC Note=No experimental confirmation available;
 CC Name=2;
 CC IsoId=Q38997-2; Sequence=VSP_009001;
 CC -1- TISSUE SPECIFICITY: Expressed in roots, shoots and leaves.
 CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. SNRP1
 CC subfamily.
 CC -1- SIMILARITY: Contains 1 UBA domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC -----
 DR EMBL; M93023; AAA2736.1; -;
 DR EMBL; X79707; CAA56146.1; -;
 DR EMBL; AC008261; AAF26165.1; -;
 DR EMBL; AY093170; AAM3169.1; -;
 DR EMBL; BT010386; AAO56829.1; -;
 DR EMBL; X94757; CAA64384.1; -;
 DR PIR; JCI146; JCI146.
 DR HSP; O63450; 1A06.
 DR InterPro; IPR001772; Kinase Cterm.
 DR InterPro; IPR000719; Prot Kinase.
 DR InterPro; IPR008271; Ser_Thr_kin_AS.
 DR InterPro; IPR002290; Ser_Thr_kinase.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF02149; KAI; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR Pfam; PF00627; UBA; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SMO0220; S_TKc; 1.

DR SMART: SM00165; UBA: 1.
 DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE: PS00111; PROTEIN KINASE DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS00300; UBA: 1.
 DR TRANSFERASE: Serine/threonine-protein kinase; ATP-binding;
 KW Phosphorylation; Alternative splicing.
 FT DOMAIN 42 294
 FT BIND 48 56
 FT ACT SITE 315 355
 FT BINDING 165 165
 FT MOD RES 198 198
 FT VARSPLIC 1 23
 FT /FTID-VSP 009001.
 SQ SEQUENCE 535 AA; 61181 MW; PFC038323FDB317 CRC64;
 Query Match 14.8%; Score 280; DB 1; Length 535;
 Best Local Similarity 28.7%; Pred. No. 2,6e-14;
 Matches 87; Conservative 46; Mismatches 118; Indels 52; Gaps 9;
 QY 80 GRAYQALHCPGTCTCKVY-----PVQALAVLEPPARLPKHKVAPTEVL 127
 DB 54 GGVKIAEHALTGKVAIKILNRKIKNMMEKVRREIKIL---RLFWAPHITRLYEVI 109
 QY 128 AGTOLLVAFETRTTH-GDMHSLVTRHRIPEPEAAVLFROMATALACHQGLVRLDK-- 184
 DB 110 EHPPTDIYVMEYVNSSELPDYIVVEKORLQEDENRNFQOIIISVEGCHNNMNVHRLKDE 169
 QY 185 -----LCRFVADRERRKLVLEN--LEDSCVLTCGDDSLMDKACPAVYGEIISRRAS 236
 DB 170 NILDSKCNVKIADFLSLNSMRDGHPLKTSK-----GSPNVAAEVVISGKL- 215
 QY 237 YEGKADVMSLSVALFTMLAGHYPPQDSEPVLLFGKIRGAYALPAGLSAPARCIVRL 296
 DB 216 YAGPEVDVWSCGIVTALLCGTLPPDENIPNLFKIKGQIYTLPSHLSPGARDLIPRYVL 275
 QY 297 REEPAERLTATGTLILPWLRODPMPLAPTRSHLMEAAQV-VPDGLGLDEAREEGDREYV 355
 DB 276 VVDPMKRVITLPIRLQHPWF-----QAHLPRIAVPPPTVQAKKIDELIAGEVI 325
 QY 356 IYV 358
 DB 326 NMG 328
 RESULT 13
 SNF1 YEAST STANDARD; PRT; 633 AA.
 ID SNF1 YEAST
 AC P06782;
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 01-JUN-1988 (Rel. 06, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Carbon catabolite derepressing protein kinase (EC 2.7.1.-)
 GN SNF1 OR CAC1 OR PAS14 OR GLC2 OR YDR477W OR DB035.-20.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OC NCBI_TaxID=4932;
 OX 11
 RN SEQUENCE FROM N.A.
 RA Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E.,
 RA Berno A., Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,
 RA Hunkeler-Smith S., Hyman R., Komp C., Lahekari D., Lew H., Lin D.,
 RA Moseedale D., Nakahara K., Namath A., Oetner P., Oh C., Fétel F.X.,
 RA Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,

RA Winant A., Yelton M., Botstein D., Davis R.W.;
 RL Submitted (AUG-1995) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RN SEQUENCE OF 274-284; 528-539 AND 622-630, AND PHOSPHORYLATION SITE.
 RP MEDLINE=94131988; PubMed=7905477;
 RA Mitchellhill K.I., Stapleton D., Gao G., House C., Michell B.,
 RA Kassis P., Witters L.A., Kemp B.E.;
 RT "Mammalian AMP-activated protein kinase shares structural and
 RT functional homology with the catalytic domain of yeast Snf1 protein
 RT kinase.";
 RL J. Biol. Chem. 269:2361-2364(1994).
 CC -FUNCTION: Essential for release from glucose repression. It
 CC interacts and has functional relationship to the regulatory
 CC protein SNF4. Interacts also with S1P1, S1P2 and GAL3. Could
 CC phosphorylates CAT8.
 CC -SUBCELLULAR LOCATION: Associated with the nuclear membrane.
 CC -SIMILARITY: Belongs to the Ser/Thr family of protein kinases. SNF1
 CC subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: M13971; AAA5058.1; -.
 DR EMBL: U33050; AAB64904.1; -.
 DR PIR: A26030; A26030.
 DR HSRP: P24941; 1HCL.
 DR GeronOnline; 140969; -.
 DR SGD: S0002885; SNF1.
 DR GO: GO:0005737; C:cytoplasm; IPI.
 DR GO: GO:0005634; C:nucleus; IPI.
 DR GO: GO:0000324; C:cytosol (sensu Fungi); IPI.
 DR GO: GO:0004675; P:SNF1/AMP-activated protein kinase activity; IDA.
 DR GO: GO:0007155; P:cell adhesion; IMP.
 DR GO: GO:0006995; P:cellular response to nitrogen starvation; IDA.
 DR GO: GO:0030447; P:filamentous growth; IMP.
 DR GO: GO:0006468; P:protein amino acid phosphorylation; IDA.
 DR GO: GO:0006109; P:regulation of carbohydrate metabolism; IGI.
 DR InterPro: IPR000719; Prot. kinase.
 DR InterPro: IPR008271; Ser_Thr_kin_AS.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR Pfam: PF00069; pkinase; 1.
 DR ProDom: PD000001; Prot. kinase; 1.
 DR SMART: SM00220; S_TKc; 1.
 DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
 KW Transferase; Serine/threonine-protein kinase; ATP-binding;
 KW Phosphorylation; Carbohydrate metabolism; Nuclear protein.
 FT DOMAIN 18 32
 FT BIND 55 306
 FT NE BIND 61 69
 FT BINDING 84 84
 FT ACT SITE 177 177
 FT MOD RES 210 210
 FT /FTID-VSP 009001.
 SQ SEQUENCE 633 AA; 72045 MW; F5C63565C986C4E3 CRC64;
 Query Match 14.5%; Score 273.5; DB 1; Length 633;
 Best Local Similarity 27.8%; Pred. No. 1e-13;
 Matches 88; Conservative 53; Mismatches 138; Indels 37; Gaps 11;
 QY 51 SPTAPDARATVAGSLGIVLLEP-EEG--GRAYQALHCPGTCTCKVYVQALAV 107
 DB 35 SNTSLNPKSSLDGAGHIGYQIVKTIGBSFGKVKLAYHTTTGQKALKLIN-KYVLAK 93
 QY 108 IE-----PYARLPKHKVAPTEVLVAGTQLVAFETRTTHGDMHSLVTRHRIPEPE 158
 DB 94 SDWGRIERERISTLRLRHPIIKLVYIKSKDEIIVIVAGNELPDIYQDRKMSGE 153

```

QY 159 AAVLFQWATALAHCHQHLVRLDKLRFVFADEKRLVLEMLDSCVLTGPDLSLMD 218
DQ 154 ARRFQOQISAVEYCHRHKIVHRLDKPENLLDHLWKLADPGL--SNIT--DGNFL 208
QY 219 KHAC--PAYQPELLSRPASVSGAADVSLGVALFTMLAGHYPODESEPVLLFGKTRG 276
DQ 209 KTSQSGSPVYAPFVTSGLT-YAGPEVYVWSCGVLLYVWLCLRLPFDDESI-PVLFKINS 267
QY 277 AVALPAGISAPARCIVCLIREPAREBLTANGILLHMLRDP--PMPLAPTRSHLMEAAQV 335
DQ 268 VYTLPRKLSRPAAGILKRLMLVPLNRTISHEIMQDDMFVLDLPEYLLP----- 316
QY 336 VPDGLGLDEAREEBCD 351
DQ 317 -PD---LKPRPEENE 328

RESULT 14
CDRL_SCHPO STANDARD; PRT; 593 AA.
AC P07334; Q9PEQ4;
DT 01-APR-1988 (Rel. 07, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mitosis inducer protein kinase cdrl (EC 2.7.1.-) (Protein kinase
DE nm1l).
GN CDRL OR NIM1 OR SPAC644.06C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_taxid=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91169281; PubMed=2004705;
RA Felletter H., Nurse P., Young P.G.;
RT "Genetic and molecular analysis of cdrl/nim1 in Schizosaccharomycetes
RT pombe.";
RL Genetics 127:309-318 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87187654; PubMed=3453113;
RA Russell P., Nurse P.;
RT "The mitotic inducer nim1+ functions in a regulatory network of
RT protein kinase homologs controlling the initiation of mitosis.";
RL Cell 49:569-576 (1987).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Baaham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Frazer A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones I., Jones M., Leather S., McDonald S., Mclean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares K., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymompres B.,
RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambut R., Purnelle B.,
RA Coffeau A., Cadieu E., Dreno S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garçon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,

```

```

RA Shpakovski G.V., Useery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomycetes pombe.";
RL Nature 415:871-880 (2002).
CC -1- FUNCTION: This protein, a dose-dependent mitotic inducer, appears
CC to function as a negative regulator of mitosis inhibitor weel by
CC phosphorylating and inactivating it.
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIM1
CC subfamily.
CC -1- CAUTION: Ref.2 sequence differs from that shown due to
CC frameshifts.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DQ EMBL; X57548; CAA40774.1; -.
DQ EMBL; M16509; AAA35317.1; ALT_FRAME.
DQ EMBL; ALJ55012; CAB90133.1; -.
DQ PIR; S16153; KIZPMN.
DQ HSSP; O63450; 1A06.
DQ GeneDB Spombe; SPAC644.06c; -.
DQ InterPro; IPR000719; Prot_Kinase.
DQ InterPro; IPR008271; Ser_Thr_kin_AS.
DQ InterPro; IPR002290; Ser_Thr_pkinase.
DQ InterPro; IPR001245; Tyr_pkinase.
DQ Pfam; PF00069; pkinase; 1.
DQ PRINTS; PR00109; TYRKINASE.
DQ ProDom; PD000001; Prot_Kinase; 1.
DQ SMART; SM00220; S_TKc; 1.
DQ PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DQ PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DQ PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DQ Transferrase; Serine/threonine-protein kinase; ATP-binding;
DQ Mitosis.
DQ DOMAIN.
DQ FT 12 258 PROTEIN KINASE.
DQ FT NP_BIND 18 26 ATP (BY SIMILARITY).
DQ FT BINDING 41 41 ATP (BY SIMILARITY).
DQ FT ACT_SITE 128 128 BY SIMILARITY.
DQ FT CONFLICT 252 252 V -> F (IN REF. 1).
DQ FT CONFLICT 570 570 A -> I (IN REF. 1).
DQ SQ SEQUENCE 593 AA; 66954 MW; DE30AE06B070F458 CRC64;

Query Match 14.3%; Score 270.5; DB 1; Length 593;
Best Local Similarity 31.5%; Pred. No. 1.6e-13;
Matches 74; Conservative 45; Mismatches 107; Indels 9; Gaps 6;

QY 85 AHCCTGCTGYTCVYVPOALAVL-PAALPPKHVAPTEVLGTOGLY-AFPTRTG 142
DQ 29 AHAATGGLAAIKIRIPYVASIGMEILMMRLRHNPILRLYVWDHGMVLALEYVPG 88
QY 143 DMSHSLVTRTRIRPEPAVLFVROMATALAHCHQHLVRLDKLRFVFADEKRLVLEMLDSCVLTGPDLSLMD 202
DQ 89 ELFHTRIRKGRPSERBAHYVLSQILDVAHCHRRFRRHRLDLENILIKVNDQ--IKI 145
QY 203 LEDSCVLTGPDLSLMDKACPA--YVGPBILSRASVSGKADVWSLGVALTMLAGHYP 260
DQ 146 ADFGAVATPEPNDSCLENY-CGSIHYLAPEIVSHK-PYGAAPADVWSCVLLYSLSNKLK 203
QY 261 FQDSFVLLFGKTRGATVAGLSAPARCIVCLIREPAREBLTANGILLHMLRDP--PMPLAPTRSHLMEAAQV 315
DQ 204 FGGQNTDVITNKRIRGAVDLPSSISSAODILHRLMDVNPSTRITRIPVESHPL 258

RESULT 15
SNP1_SCHPO STANDARD; PRT; 576 AA.
AC Q74536;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)

```

DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE SNF1-like protein kinase (BC 2.7.1.-).
GN SPC74.03C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxId=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21849401; PubMed=11859360;
RA Wood V., Galliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Pat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean U.,
RA Mooney P., Mould S., Mungall K., Murphy L., Niblett D., Odeil C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben U., Grymoprez B.,
RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Leleure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Szpakowski G.V., Ussey D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. SNF1
subfamily.
CC
CC -!- SIMILARITY: Contains 1 UBA domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: AL031543; CAA20833.1; -.
CC PTR: T41587; T41587.
CC
CC HSSP: Q63450; 1A06.
CC GeneDB: SPombe, SPC74.03C; -.
CC InterPro: IPR000719; Prot_kinase.
CC InterPro: IPR008271; Ser_thr_kin_AS.
CC InterPro: IPR002290; Ser_thr_kinase.
CC InterPro: IPR000449; UBA_domain.
CC Pfam: PF00069; kinase_1.
CC ProDom: PD000001; Prot_kinase; 1.
CC SMART: SM00220; S_TKc; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
CC PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
CC DR PROSITE: PS50030; UBA_1.
CC KW transferase; Serine/threonine-protein kinase; ATP-binding.
CC FT DOMAIN 34 285 PROTEIN_KINASE.
CC FT NP_BIND 304 345 UBA.
CC FT BINDING 40 48 ATP (BY SIMILARITY).
CC FT ACT_SITE 63 63 ATP (BY SIMILARITY).
CC FT ACT_SITE 156 156 BY SIMILARITY.
CC SQ SEQUENCE 576 AA; 65996 MW; E5857E8F171E7B50 CRC64;

Query Match 14.3%; Score 270; DB 1; Length 576;
Beet Local Similarity 25.8%; Pred. No. 1.7e-13;
Matches 83; Conservative 54; Mismatches 133; Indels 52; Gaps 11;
QY PPTAPDRATAVATASRLGPVYLLEP-EEG--GRAVQALHCPGTGEYTKVVP-----100
DB 22 PPEA-----ISKRHITPIIIRRTSGSGKVLALHYHTQGVALKFISRLQALKSSD 74
QY 101 ---VQELAVLEPYARLPKHKVAPREVLATGQLLYAFPTTHGDMSLVTRTHRIPE 156
DB 75 MEMRVEREIS---YTKLRHPHITIKLDVITTPFDIVWVLEVAGGELFDYIVKKRMTE 130
QY 157 PEAAVLPRQMATLALAHQHGHGIVLBDILKCFVFPADREKRVLENLEDSCVLTGPDSTL 216
DB 131 DEGRFPFOOICALIYCHRHKIVHBDLKPEMLLDNNLVKIADEGL--SNMT--DGN 185
QY 217 WDKHAC--PAYVGPETLSRSASYSKADVMSLGVAFETMLAGHYFQDSEPVLLFGKIR 274
DB 186 FLKTCGSPNVAPEVINGKL-YAGPEVDVMSGIVLYVMLVGRLPFDDEFIPNLFKYN 244
QY 275 RGAVYLPAGLSAPARCLVRCLLRREPARLTATGILLHPWRQDPMPLAPTRSHLWEAAQ 334
DB 245 SCVYVMPDPLSPGQSLIRMTIVADPMQRIITQRIKRPWFNVN-----288
QY 335 VVPDGL-GLDERAREEGPREVY 355
DB 289 -LPDYLRPMEEVQSGSYADSRIV 309

Search completed: August 24, 2004, 18:53:20
Job time : 14 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 18:49:06 ; Search time 45 Seconds
(without alignments)
2510.123 Million cell updates/sec

Title: US-10-070-337-5
Perfect score: 1892
Sequence: 1 MRATPLAAPSISRRKRL.....GLGDEAREBEGDEVLVYG 358

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP RODENT:*
12: SP_VIRUS:*
13: SP_VERTebrATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	766.5	40.5	343	13	07ZZY2
2	759.5	40.1	343	6	0864R4
3	756	40.0	343	11	08K017
4	754.5	39.9	343	6	Q28283
5	753.5	39.8	343	4	Q92519
6	753	39.8	343	11	08K4K3
7	742.5	39.2	364	11	09EGT6
8	737.5	39.0	372	4	096R08
9	735.5	38.9	372	4	09H2V8
10	735.5	38.9	372	11	091W04
11	722.5	38.2	372	11	08K4K4
12	717	37.9	344	13	07ZZZ7
13	657	34.7	257	11	08R2V8
14	608.5	32.2	224	4	015180
15	432	22.8	484	5	09V3Z1
16	341.5	18.0	443	10	07X1W5

17	319	16.9	218	11	08BFS7	08BFS7 mus musculu
18	319	16.9	379	11	08BUT9	08BUT9 mus musculu
19	319	16.9	388	11	08BZX3	08BZX3 mus musculu
20	318	16.8	461	10	091GV5	091GV5 oryza sativ
21	313.5	16.6	706	3	08J2N0	08J2N0 fusarium ox
22	305	16.1	474	10	094DR7	094DR7 oryza sativ
23	303.5	16.0	446	4	096CV1	096CV1 homo sapien
24	299	15.8	914	5	019469	019469 caenorhabdi
25	298.5	15.8	446	10	08H5S1	08H5S1 oryza sativ
26	298	15.8	504	11	08K0J7	08K0J7 mus musculu
27	296	15.6	880	3	09Y880	09Y880 coxliobol
28	295	15.6	671	3	096M17	096M17 trichoderma
29	293	15.5	777	3	0872H0	0872H0 neurospora
30	291.5	15.4	443	13	07T0B0	07T0B0 xenopus lae
31	291.5	15.4	651	13	091821	091821 xenopus lae
32	290	15.3	602	3	P87209	P87209 kluyveromyc
33	288.5	15.2	698	5	061298	061298 halocynthia
34	288	15.2	688	5	095UP4	095UP4 ancylostoma
35	287.5	15.2	676	13	07ZZN5	07ZZN5 brachydanto
36	287	15.2	477	10	091W06	091W06 oryza sativ
37	286.5	15.1	676	13	07ZU72	07ZU72 brachydanto
38	284	15.0	514	10	09M726	09M726 lycopersico
39	283	15.0	504	10	P93113	P93113 cucumis sat
40	283	15.0	512	10	P92958	P92958 arabidopsis
41	282	14.9	438	10	081K24	081K24 glycine max
42	282	14.9	512	10	P92968	P92968 arabidopsis
43	281.5	14.9	419	4	09H7H6	09H7H6 homo sapien
44	280.5	14.8	643	11	061804	061804 mus musculu
45	280.5	14.8	643	11	061846	061846 mus musculu

ALIGNMENTS

RESULT 1

ID 07ZZY2 PRELIMINARY; PRT; 343 AA.

AC 07ZZY2; 01-JUN-2003 (TREMBLrel. 24, Created)

DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)

DR 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

Query Match 40.5%; Score 766.5; DB 13; Length 343;

Best Local Similarity 47.6%; Pred. No. 2.9e-57;

Matches 168; Conservative 50; Mismatches 114; Indels 21; Gaps 6;

QY 2 MRATPLAAPSISRRKRL...LDTERRPQKARSGPQPRLLPCLLPSPPTAPPRATA 61

DB 5 RSPITITARYGRGRNKTQDEE-LSSIRSLP-----PQSGFSPNUGSPSPPPPPPNUSHC 57

QY 62 VATA⁸RLG¹YV¹LE¹PE¹BE¹GR¹A¹Q¹ALH¹CP¹GT¹CT¹KY¹--Q¹EA¹LV¹LE¹PA¹RI¹PK¹ 118

D¹b 58 V---SC¹IG¹KT¹LE¹PE¹GD¹H¹V¹RA¹V¹ALH¹SG¹EE¹LV¹CC¹VF¹D¹GCY¹O¹EL¹LA---PC¹CP¹PA¹AK¹ 111

QY 119 HV¹AR¹PE¹VL¹AG¹T¹OL¹LY¹AF¹PT¹TH¹GD¹MS¹LV¹TR¹HR¹I¹PE¹BA¹VL¹PR¹OM¹AT¹ALH¹CH¹Q¹GL¹ 178

D¹b 112 NI¹Q¹IT¹EL¹IL¹GT¹KA¹V¹FE¹RS¹Y¹GD¹MS¹FL¹RT¹CK¹L¹KE¹BA¹VL¹FP¹Q¹MS¹VA¹HC¹HD¹GL¹ 171

QY 179 VL¹RD¹KL¹CR¹FA¹DA¹DE¹R¹CK¹VL¹EN¹ED¹SC¹VL¹GP¹DD¹SL¹MD¹KA¹CP¹AV¹GC¹PE¹IL¹SR¹AS¹ 238

D¹b 172 VL¹RD¹KL¹KR¹FL¹K¹DE¹RI¹K¹V¹KL¹ES¹LE¹DA¹IL¹IG¹ND¹DS¹L¹DK¹CP¹AV¹SP¹EI¹INT¹GS¹YS¹ 231

QY 239 GK¹AA¹DV¹MS¹L¹Q¹AL¹FT¹ML¹AG¹HY¹PO¹DE¹PA¹VL¹L¹FG¹K¹IR¹GA¹Y¹AL¹P¹AG¹AS¹AP¹AC¹LV¹RC¹L¹IR¹ 298

D¹b 232 GK¹AA¹DV¹MS¹L¹Q¹AL¹FT¹ML¹VR¹YP¹FD¹HE¹PE¹SL¹SK¹IR¹Q¹GN¹IP¹ET¹LS¹PA¹K¹CI¹RI¹EL¹IR¹ 291

QY 299 E¹PA¹EL¹TA¹IG¹IL¹HP¹ML¹RQ¹DP¹ML¹AP¹TR¹SH¹LM¹EA¹Q¹V¹VD¹GL¹GD¹EA¹RE¹EC¹D¹ 351

D¹b 282 E¹PE¹SE¹LT¹Q¹EL¹IL¹HP¹FT¹SD---FN¹VS¹NG¹GA¹EA¹VS¹DO¹LP¹VD¹NME¹ED¹L¹ 339

	RESULT	2
Q864R4	P0864R4	PRELIMINARY; PRT; 343 AA.
AC	Q864R4;	
DT	01-JUN-2003 (TrEMBLrel. 24, Created)	
DT	01-JUN-2003 (TrEMBLrel. 24, Last sequence update)	
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)	
DE	TBR2 protein.	
OS	Bos taurus (bovine).	
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea,	
OX	Bovidae; Bovinae; Bos.	
RN	NCBI_Taxid=9913;	
RP	[1]	
RA	SEQUENCE FROM N.A.	
RL	Shan Y.X., Yu L.;	
RU	Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.	
DR	EMBL; AY247741; AAP04410.1; -	
DR	GO; GO:0005524; F:ATP binding; IEA.	
DR	GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.	
DR	GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.	
DR	GO; GO:0006468; P:protein amino acid phosphorylation; IEA.	
DR	InferPro; IPR000719; Prot_kinase.	
DR	InferPro; IPR002290; Ser_thr_kinase.	
DR	Pfam; PF00069; pkinase; Tyr_pkinase.	
DR	Prodrom; PD000001; Prot_kinase; 1.	
DR	SMART; SM00220; S_TKc; 1.	
DR	SMART; SM00219; TYR_KC; 1.	
DR	PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.	
DQ	SEQUENCE_343_AA; 38791 MW; 070416915A4145AC CRC64;	

Query March	40.1%;	Score 759.5;	DB 6;	Length 343;
Best Local Similarity	47.6%;	Pred. No. 1.1e-56;		
Matches 168;	Conservative 50;	Mismatches 114;	Indels 21;	Gaps 6;

QY 2 R A T E L A A P A G S L S R K K K R E L D N D I T E P V Q K R A S G Q P R L P C L L P L S E P T A D R A T A 61
 Db 5 R S I T T I A R Y S R N K K T O D F E E - L S I S I S A E - - - - P S Q S F P M I G S P S P E T N L S H C 57
 QY 62 V A T A S R L G P Y V L L E P E E G G R A V O A L H C P T G T E Y T C K V P V - - Q E A L A V L E P A R L P R H X 118
 Db 58 V - - - S C I G K Y L L L E P L B E D H V P R A V H L S G E L V A K V C I D I S Y Q S L A - - - P C F C L S A H S 111
 QY 119 H V A R P T E V L A G T O L L I Y A F F T R T H G D M S L V A T R R H R I P E E A A V L E R O M A T A L A H C H Q G L 178
 Db 112 N I N O I T E I I L I G E T A Y A V F F E R S Y G D M S F V A T C K L R E E A A R L P Y O I A S A V A H C D G L 171
 QY 179 V L N D I K L C R F E A D R E R K C V L N E L D S C V L T G P D S L W D K A C P A Y V G P E L L S R A S Y 238
 Db 172 V L N D L K R R K T F L K O E E R T R V L S L E A V I L R G D D S L S D R G C P A Y V S P E L L N N G S Y 231

QY 239 GKADVMSLGAALFTMLAGHYRPODSEPVLLFGKIRGAYALPAGISAPARCLVRCLLR 258
Db 232 GAAAVMSLGNVLYMLVGRYRFOIEBSLSFSKIRRGQFNIPETLSPKCLIRILRR 251
QY 239 EPAERLTATGILLHPMLRQDPMPLAFTSRSHWEAAQVVDGIGLDEAREEGD 351
Db 232 EPSEELTSGEILIDHEMFSTD-----FSVNSGSGAAGVSDOLVEDVNMETLL 339

```

RESULT 3
Q8K017
ID Q8K017 PRELIMINARY; PRT; 343 AA.
AC Q8K017;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
DE TRR-2 (CSFW ORF protein homolog) .
GN AWJ19517.
OS Mus musculus (Mouse) .
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
NN [2]

```

AP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Aorta, Vein, Cerebellum, and Retina;
RX MEDLINE=223545683; PubMed=12465651;
RA THE PANTOM Consortium.
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; BC034338; AAH34338.1; -
DR EMBL; AK044747; BAC32063.1; -
DR EMBL; AK080064; BAC37820.1; -
DR EMBL; AK082329; BAC38467.1; -
DR MGI; MGI:2145021; AW319517.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004673; P:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004173; P:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0004688; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR007019; Prot Kinase.
DR InterPro; IPR002290; Ser Thr Kinase.
DR InterPro; IPR01245; Tyr Kinase.
DR Pfam; PF00069; pkinase; 1.
DR Prodom; PD000001; Prot Kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Transferase.
QO SEQUENCE 343 AA; 38772 MW; 9418B7AC19FCC23F CRC64;

Query March Similarity	40.0%	Score 756;	DB 11;	Length 343;
Best Local Similarity	46.9%	Pred. No. 2.3e-56;		
Matches 164; Conservative	48;	Mismatches 108;	Indels 30;	Gaps 6;

[illegible]

QY 179 VLRLDKLGRFVADDERKKLVLENIEDSCVLTGPDSDIMDKACAPAYGPELTLSSRASYS 238
 DB 172 VLRLDKLKRKFKEDEERTVKLESLIEDAYILRGDDSDISDKHGCAPAYSPETLNTSGSYS 231
 QY 239 GKADVWSLGVALLFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVACLR 298
 DB 232 GKADVWSLGVALLFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVACLR 291
 QY 299 EPARLRTATGILLHPMLRQDPMPLAPTRSHLWEAAQVVPDGLDEAREE 348
 DB 292 EPARLRTATGILLHPMLRQDPMPLAPTRSHLWEAAQVVPDGLDEAREE 327

RESULT 4

Q28283 PRELIMINARY; PRT: 343 AA.
 ID 028283;
 AC 028283;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE C5FW ORF protein.
 GN C5FW ORF.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thyroid;
 RX MEDLINE=97067069; PubMed=8910471;
 RA Wilkin F., Savonet V., Radulescu A., Petermans J., Dumont J.E.,
 RA Maenhaut C.;
 RT "Identification and Characterization of Novel Genes Modulated in the
 RT Thyroid of Dogs Treated with Methimazole and Propylthiouracil."
 RL J Biol. Chem. 271:28451-28457(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thyroid;
 RX MEDLINE=98000262; PubMed=9342215;
 RA Wilkin F., Suarez-Huerta N., Robaye B., Peetermans J., Libert F.,
 RA Dumont J.E., Maenhaut C.;
 RT "Characterization of a phosphoprotein whose mRNA is regulated by the
 RT mitogenic pathways in dog thyroid cells."
 RL Eur. J. Biochem. 248:660-669(1997).
 DR EMBL; X99144; CAA67581.1; -.
 DR HSSP; O63450; 1A06.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004672; F:protein kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR00719; Prot kinase.
 DR Pfam; PF00069; pkinae; 1.
 DR ProDom; PD000001; Prot kinase; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW ATP-binding; Transferase.
 SQ SEQUENCE 343 AA; 38786 MW; BF8D1300DACB84FA CRC64;

Query Match 39.9%; Score 754.5; DB 6; Length 343;

Best Local Similarity 47.3%; Pred. No. 3e-56; Mismatches 115; Indels 21; Gaps 6;

QY 2 RATPLAAPAGSLSRKKRLDNLDERPVQKRSQPPRLPCLPLSPPTAPRARATA 61
 DB 5 RSTPLTIARVGRSRNKTQDFEE-LSISIRSAE-----PSQSPSPNLSPPSPPTPLSHC 57
 QY 62 VATAARLGPVYLLIEEGRAYQALHCPGTETCTKVYV---QALAVLEBYARLPK 118
 DB 58 V---SCIGKYLLEPGLGHPRAVHMSGEELVCKVPDISCYQSLA---PCFCLSAHS 111
 QY 119 HVARPEVLAQTOLYAFTTRTHGDMHSLVTRHRIPEPEAAVLFQKATALAHQHGSL 178
 DB 112 NINQITTEIILGRTKAYVFERSYGDMSFVCTCKLREBEARLFTQIASAVAHCHDGL 171

QY 179 VLRLDKLGRFVADDERKKLVLENIEDSCVLTGPDSDIMDKACAPAYGPELTLSSRASYS 238
 DB 172 VLRLDKLKRKFKEDEERTVKLESLIEDAYILRGDDSDISDKHGCAPAYSPETLNTSGSYS 231
 QY 239 GKADVWSLGVALLFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVACLR 298
 DB 232 GKADVWSLGVALLFTMLAGHYPFQDSEPVLLFGKIRRGAYALPAGLSAPARCLVACLR 291
 QY 299 EPARLRTATGILLHPMLRQDPMPLAPTRSHLWEAAQVVPDGLDEAREE 351
 DB 292 EPARLRTATGILLHPMLRQDPMPLAPTRSHLWEAAQVVPDGLDEAREE 339

RESULT 5

Q92519 PRELIMINARY; PRT: 343 AA.
 ID 092519;
 AC 092519;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE GS395 (TRB2 protein).
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cancellous bone;
 RA Ohno I., Hashimoto J., Takaoka K., Ochi T., Okubo K., Matsubara K.;
 RL Submitted (AUG-1996) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=uterus;
 RA Strausberg R.;
 RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Shan Y.X., Yu L.;
 RL Submitted (FEB-2003) to the EMBL/Genbank/DBJ databases.
 DR EMBL; D87119; BAA13250.1; -.
 DR EMBL; BC002637; AA02637.1; -.
 DR EMBL; AY245544; AA089231.1; -.
 DR HSSP; Q63450; 1A06.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004672; F:protein kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR00719; Prot kinase.
 DR Pfam; PF00069; pkinae; 1.
 DR ProDom; PD000001; Prot kinase; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW ATP-binding; Transferase.
 SQ SEQUENCE 343 AA; 38800 MW; BF8B7366DACB84FA CRC64;

Query Match 39.8%; Score 753.5; DB 4; Length 343;

Best Local Similarity 47.3%; Pred. No. 3.7e-56; Mismatches 115; Indels 21; Gaps 6;

QY 2 RATPLAAPAGSLSRKKRLDNLDERPVQKRSQPPRLPCLPLSPPTAPRARATA 61
 DB 5 RSTPLTIARVGRSRNKTQDFEE-LSISIRSAE-----PSQSPSPNLSPPSPPTPLSHC 57
 QY 62 VATAARLGPVYLLIEEGRAYQALHCPGTETCTKVYV---QALAVLEBYARLPK 118
 DB 58 V---SCIGKYLLEPGLGHPRAVHMSGEELVCKVPDISCYQSLA---PCFCLSAHS 111
 QY 119 HVARPEVLAQTOLYAFTTRTHGDMHSLVTRHRIPEPEAAVLFQKATALAHQHGSL 178
 DB 112 NINQITTEIILGRTKAYVFERSYGDMSFVCTCKLREBEARLFTQIASAVAHCHDGL 171
 QY 179 VLRLDKLGRFVADDERKKLVLENIEDSCVLTGPDSDIMDKACAPAYGPELTLSSRASYS 238

DB 172 VLRLDLKLRKFIKDEERTVYKLESLDPAVILRGDDSLSDKGGCPAVVSPFELINTSGSYS 231

QY 239 GKADWVMSLGVALLFTMLAGHYPPQDSEPVLLFGKIRRGAYALPAGLSAPRCCLVRCLLR 298

DB 232 GKADWVMSLGVALLFTMLAGHYPPHDIPESSLSFKIRRGQFNIPETLSPPKACILRSLRR 291

QY 299 EPAERLTATGILHLPWLRODPMPLAFTRSHLWEAAQVVPDGLGIDEAREEGD 351

DB 292 EPERLTSOELIDHPWFSTD-----FSVNSAIGAKEVSLQVLPDVAMEEND 339

RESULT 6

Q8K4K3 PRELIMINARY; PRT; 343 AA.

AC 08K4K3;

DB 01-OCT-2002 (T-EMBLrel. 22, Created)

DT 01-OCT-2002 (T-EMBLrel. 22, last sequence update)

DE 01-OCT-2003 (T-EMBLrel. 25, last annotation update)

GN TRB-2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Kiss-Toth E., Dempsey C., Jozsa V., Caunt J., O'Leary K.M., Bagnsaff S.M., Wyllie D.H., Harte M., O'Neill L.A.U., Owarstrom E.E., Dower S.K.;

RT "Mammalian homologs of Drosophila tritbbles (trtb) control mitogen activated protein kinase signaling.";

RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF358867; AM45477.1; -

DR MCD; MGJ:2345021; AM319517.

DR GO; GO:000524; F:ATP binding; IEA.

DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.

DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.

DR GO; GO:0016740; F:transferase activity; IEA.

DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.

DR InterPro; IPR000719; Prot_kinase.

DR InterPro; IPR002290; Ser_Thr_kinase.

DR InterPro; IPR001245; Tyr_kinase.

DR Pfam; PF00069; pkinase; 1.

DR ProDom; PD000001; Prot_kinase; 1.

DR SMART; SM00220; S_TKc; 1.

DR SMART; SM00219; TYKc; 1.

DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

DR ATP-binding; Transferase.

KW SEQUENCE 343 AA; 38758 MW; 0B3965B8B2087D74 CRC64;

Query Match 39.8%; Score 753; DB 11; Length 343;

Best Local Similarity 46.6%; Pred. No. 4.1e-56;

Matches 163; Conservative 49; Mismatches 108; Indels 30; Gaps 6;

QY 2 RATPLAAPAGSLSRKKRLDNLDTFRVYOKRARGPOPLPCLLPSPPTAPRATA 61

DB 5 RSTPLTIARYGSRNKTQDFEE-LSSTRAB-----PSQSPFNCGSPSPETPLNSHC 57

QY 62 VATASLGGVYLLPEEGRAYOALHCPGTGYTCVYV--QDALVLEPYARLPK 118

DB 58 V---SCITGKYLLLEPDEGHVFAVHSHGEEIVCVFETSCQESIA--PCFCISAS 111

QY 119 HVAPREVIAGTQLTAFFTRTHGMSLVTRTHRIPEPEAVLPFQOMATALACHQHQL 178

DB 112 NINQITREILIGETKAVYFFERSIGDHSYVTCCKLRSEBAARLFYQISAVAHCHDGL 171

QY 179 VLRLDLKLRKFIKDEERTVYKLESLDPAVILRGDDSLSDKGGCPAVVSPFELINTSGSYS 238

DB 172 VLRLDLKLRKFIKDEERTVYKLESLDPAVILRGDDSLSDKGGCPAVVSPFELINTSGSYS 231

QY 239 GKADWVMSLGVALLFTMLAGHYPPQDSEPVLLFGKIRRGAYALPAGLSAPRCCLVRCLLR 298

DB 232 GKADWVMSLGVALLFTMLAGHYPPHDIPESSLSFKIRRGQFNIPETLSPPKACILRSLRR 291

QY 299 EPAERLTATGILHLPWLRODPMPLAFTRSHLWEAAQVVPDGLGIDEAREEGD 348

DB 292 EPERLTSOELIDHPWFSTD-----FSVNSGFGAKBACQD 327

RESULT 7

Q9E0L6 PRELIMINARY; PRT; 364 AA.

AC 09E0L6;

DB 01-MAR-2001 (T-EMBLrel. 16, Created)

DT 01-MAR-2001 (T-EMBLrel. 16, last sequence update)

DE 01-MAR-2003 (T-EMBLrel. 25, last annotation update)

GN G-protein-coupled receptor induced protein G1G2 (Fragment).

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RA STRAIN=Sprague-Dawley; TISSUE=Brain; Mayhaus M., von der Kammer H., Klaudiny J., Albrecht C., Hoffmann B., Nitsch R.M.;

RT "Identification of a novel nuclear factor G1G2, as an ml-acetylcholine receptor-induced gene.";

RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF205438; AAG3664.1; -

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0004672; F:protein kinase activity; IEA.

DR GO; GO:0004872; F:receptor activity; IEA.

DR GO; GO:0016740; F:transferase activity; IEA.

DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.

DR InterPro; IPR000719; Prot_kinase.

DR Pfam; PF00069; pkinase; 2.

DR ProDom; PD000001; Prot_kinase; 1.

DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

KW ATP-binding; Receptor; Transferase.

FT NON_TER 1

SEQUENCE 364 AA; 40377 MW; 8F9167FB76DFCD37 CRC64;

Query Match 39.2%; Score 742.5; DB 11; Length 364;

Best Local Similarity 45.6%; Pred. No. 3.5e-55;

Matches 161; Conservative 48; Mismatches 119; Indels 25; Gaps 5;

QY 2 RATPLAAPAGSLSRKKRLDNLDTFRVYOKRARGPOPLPCLLP--ISPPTAP-- 56

DB 10 RGPALFFPARAGTPAKRL-----LDTDDAAVAAC---PRLSECSNPPDYLSPPGSPCS 61

QY 57 -----DRATVATASRLGVPYLLPEEGRAYOALHCPGTGYTCVYVQDAL 105

DB 62 PPPPSAQTGSGCVSAPGPSRIADTLPLAERHVSALCIHTREIRCKVFPKHYQ 121

QY 106 AVLEPYARLPKHHVAPREVIAGTQLTAFFTRTHGMSLVTRTHRIPEPEAVLPFQ 165

DB 122 DKIRPYTQLPSRNTIGIVEILGSKAVYFFEXKPGDHSYVRSKRLREBAARLFQ 181

QY 166 MATALAHCHQHGLVLDLKLCPVPADEKRLVLENLDSCVLTGPDSDLMDKACPY 225

DB 182 IYSAVAHCHQSAIVIGDELKRFVSTERTQLRLSESLDTIMIKEDALSDKHCCPY 241

QY 226 VGEPIISSPRASYSGKAADVMSLGVALLFTMLAGHYPPQDSEPVLLFGKIRRGAYALPAGS 285

DB 242 VSEPIINTTIGTSGKAADVMSLGVALLFTMLAGHYPPHDSDBALFSKIRRGQFCLEHVS 301

QY 286 APARCLVRCLLRPEERLTATGILHLPWLRODPMPLAFTRSHLWEAAQVDP 338

DB 302 PARACILRLRLRPEERLTATGILHLPWFVEYVLE--GVDSVGYGSDIIVE 353

RESULT 8

Q96R08 PRELIMINARY; PRT; 372 AA.

ID 096R08

AC Q96RU8; 01-DEC-2001 (T-EMBLrel. 19, Created)
 DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
 DE SKIP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kiss-Toth E., Wylie D.H., Ovarstrom E.E., Dower S.K.;
 RT "Identification of pro-inflammatory cytokine signalling network
 components by transcription expression screening."
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF250310; AAKS8174.1; -
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004672; F:protein kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000719; Prot_kinase.
 DR Pfam; PF00069; pkinase; 2.
 DR Prodom; PD000001; Prot_kinase; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR ATP-binding; Transferase.
 KW SEQUENCE 372 AA; 40980 MW; 3E2B5C97A4F98FDB CRC64;

Query Match 39.0%; Score 737.5; DB 4; Length 372;
 Best Local Similarity 46.2%; Pred. No. 9.5e-55;
 Matches 160; Conservative 49; Mismatches 126; Indels 11; Gaps 6;

QY 2 RATPLAAPA-GSLSRKKRLLEDD--NLDTERPVQKRAASGP---QRLPFCILPLSPPTA 55
 DB 17 RGPALLFPATKRCVPKRLDDADDAVAACPRLSECCSPDYLSPPGSPC-SPPPPPA 75
 QY 56 PDRATAVATA--SRIGPVYLLPEEGGRAYQALHCPGTETXCKYVPQBALAVLEPYA 112
 DB 76 PGAGGSGSAPGPSRIADYLLPLAREHVSALCIHGRRLRCVFPPIKHQDKIRPYI 135
 QY 113 RLPRKHVARPEVLVAGTQLYAFRTTHGDMHSLVTRHRIPEPEAAVLFROMATALAH 172
 DB 136 QLPSSHSNITGIVEVILGETKAYVFPEKSGFDMHYSVRSRKLREBEAARLFQIYSAVAH 195
 QY 173 CHQHLVLRDLKLCGFVADDERKTLVLENDSCVLGPPDLSMDKACPAVYGPETLS 232
 DB 196 CHQSAIVIGDLKRFVFTSTERTQLRLESLETHIMKGEDDASDKKGCPCAVYSPETLN 255
 QY 233 SRASYSGRADVWSIGVALFTMLAGHYVPDSEPVLLFGKIRGAYALPAGISAPARCLV 292
 DB 256 TTGTYSGRADVWSIGVALYTLVGRYFPFHSDDPSALFSKIRRGQFCIPHHISPARCLT 315
 QY 293 RCLLRPEFAERLTATGILLHFWLRQDPMPPLAPTSRLHMAAQQVVD 338
 DB 316 RSLLRREPSERLTAPEILHWPESVLEP-GYIDSEIGTSQIYVE 360

RESULT 9

Q9H2Y8 PRELIMINARY; PRT; 372 AA.
 AC Q9H2Y8;
 DT 01-MAR-2001 (T-EMBLrel. 16, Created)
 DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
 DE G-Protein-coupled receptor induced protein G1G2.
 GN G1G2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Mayhans M., von der Kammer H., Klaudiny J., Albrecht C., Hoffmann B.,

RA Nitsch R.M.;
 RT "Identification of a novel nuclear factor G1G2, as an ml-acetylcholine
 receptor-induced gene."
 RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF205437; AAG35663.1; -
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004672; F:protein kinase activity; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000719; Prot_kinase.
 DR Pfam; PF00069; pkinase; 2.
 DR Prodom; PD000001; Prot_kinase; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR ATP-binding; Receptor; Transferase.
 KW SEQUENCE 372 AA; 41008 MW; 5F54BE592AB1365B CRC64;

Query Match 38.9%; Score 735.5; DB 4; Length 372;
 Best Local Similarity 46.2%; Pred. No. 1.4e-54;
 Matches 160; Conservative 48; Mismatches 127; Indels 11; Gaps 6;

QY 2 RATPLAAPA-GSLSRKKRLLEDD--NLDTERPVQKRAASGP---QRLPFCILPLSPPTA 55
 DB 17 RGPALLFPATKRCVPKRLDDADDAVAACPRLSECCSPDYLSPPGSPC-SPPPPPA 75
 QY 56 PDRATAVATA--SRIGPVYLLPEEGGRAYQALHCPGTETXCKYVPQBALAVLEPYA 112
 DB 76 PGAGGSGSAPGPSRIADYLLPLAREHVSALCIHGRRLRCVFPPIKHQDKIRPYI 135
 QY 113 RLPRKHVARPEVLVAGTQLYAFRTTHGDMHSLVTRHRIPEPEAAVLFROMATALAH 172
 DB 136 QLPSSHSNITGIVEVILGETKAYVFPEKSGFDMHYSVRSRKLREBEAARLFQIYSAVAH 195
 QY 173 CHQHLVLRDLKLCGFVADDERKTLVLENDSCVLGPPDLSMDKACPAVYGPETLS 232
 DB 196 CHQSAIVIGDLKRFVFTSTERTQLRLESLETHIMKGEDDASDKKGCPCAVYSPETLN 255
 QY 233 SRASYSGRADVWSIGVALFTMLAGHYVPDSEPVLLFGKIRGAYALPAGISAPARCLV 292
 DB 256 TTGTYSGRADVWSIGVALYTLVGRYFPFHSDDPSALFSKIRRGQFCIPHHISPARCLT 315
 QY 293 RCLLRPEFAERLTATGILLHFWLRQDPMPPLAPTSRLHMAAQQVVD 338
 DB 316 RSLLRREPSERLTAPEILHWPESVLEP-GYIDSEIGTSQIYVE 360

RESULT 10

Q91W04 PRELIMINARY; PRT; 372 AA.
 AC Q91W04;
 DT 01-DEC-2001 (T-EMBLrel. 19, Created)
 DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
 DE Similar to phosphoprotein regulated by mitogenic pathways.
 GN TRB1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Breast tumor;
 RA Strausberg R.;
 RL submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Skin;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).

DR EMBL; BC006800; AA006800.1; -.
 DR EMBL; AK028626; BAC26038.1; -.
 DR MGI; MGI:2443397; Trb1.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004672; F:protein kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000719; Prot_kinase.
 DR Pfam; PF00069; kinase; 2.
 DR PROSITE; PS00001; Prot_kinase; 1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 DR ATP-binding; Transferase.
 DR SEQUENCE 372 AA; 41281 MW; AD29B4B640B462 CRC64;
 Query Match 38.9%; Score 735.5; DB 11; Length 372;
 Best Local Similarity 46.2%; Pred. No. 1.4e-54;
 Matches 163; Conservative 46; Mismatches 119; Indels 25; Gaps 6;
 QY 2 RATPLAAGSLSRKKRLTDNLDTERRVQKARSGPRLPCLP---LSPPTAP-- 56
 DB 17 RGPGLFPAAGTTPAKRL-----LPTD--DAGAAVAAKCPRLSECSPPDYLSPPGSPCS 68
 QY 57 -----DATAVATASRLGP-----YVLLPEBEGRAVQALHCPGTGYTCVVVQVAL 105
 DB 69 PQPPTGTGTGSCVSSPSPRIADYLLPLAERHVSRLCIHGRRLCKEPIKHYQ 128
 QY 106 AVLEPYARLPKHVARPTEVLAQTQLYAFTTRHGMHSIVTRHRIPEEAAVLRQ 165
 DB 129 DKIRFYIQLPSSHNTIGIVELVLSKAVVFEEKFGDMHSYVRSKRLREEARLFRQ 188
 QY 166 MATALAHCHQGLVLRDLKCRFVADREKKLVLENLSDSCVLTPGDDSLMDKACPAY 225
 DB 189 IVSAVAHCHQSAIVLGDILKRFVSTERRTQRLGSLDTHIIKGEDDALSDKHGCPAY 248
 QY 226 VGPETLSRSASGKAADVWSLGVALLFTMLAGHYFODSEPVLLRGKIRGAYALPAGLS 285
 DB 249 VSPETLNTGTYSGKAADVWSLGVALLFTMLAGHYFODSEPVLLRGKIRGAYALPAGLS 308
 QY 286 APARCLVCLRLRREPARLITAGILHPMLRQDPMPLATRSHLWEAAQVVD 338
 DB 309 PPARCLIRLSLRREPSERLTAQIILHPMFEVLEP-GYVSEIGTSQDIVE 360
 RESULT 11
 Q8K4K4 PRELIMINARY; PRT; 372 AA.
 AC Q8K4K4;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, last annotation update)
 DE TRB-1.
 GN TRB1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kiss-Toch E., Dempsey C., Jozsa V., Cant J., Oxley K.M.,
 RA Bagstaff S.M., Wyllie D.H., Harre M., O'Neill L.A.D., Qvarnstrom E.E.,
 RA Power S.K.;
 RT "Mammalian homologs of Drosophila trilles (http://control.mitogen
 RT activated protein kinase signaling";
 RT Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AF358866; AAM45478.1; -.
 DR MGI; MGI:2443397; Trb1.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR Pfam; PF00069; kinase; 2.

DR PRODOM; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 DR ATP-binding; Transferase.
 DR SEQUENCE 372 AA; 41282 MW; 3A3DE82B46CD907F CRC64;
 Query Match 38.2%; Score 722.5; DB 11; Length 372;
 Best Local Similarity 45.6%; Pred. No. 1.9e-53;
 Matches 161; Conservative 46; Mismatches 121; Indels 25; Gaps 6;
 QY 2 RATPLAAGSLSRKKRLTDNLDTERRVQKARSGPRLPCLP---LSPPTAP-- 56
 DB 17 RGPGLFPAAGTTPAKRL-----LPTD--DAGAAVAAKCPRLSECSPPDYLSPPGSPCS 68
 QY 57 -----DATAVATASRLGP-----YVLLPEBEGRAVQALHCPGTGYTCVVVQVAL 105
 DB 69 PQPPTGTGTGSCVSSPSPRIADYLLPLAERHVSRLCIHGRRLCKEPIKHYQ 128
 QY 106 AVLEPYARLPKHVARPTEVLAQTQLYAFTTRHGMHSIVTRHRIPEEAAVLRQ 165
 DB 129 DKIRFYIQLPSSHNTIGIVELVLSKAVVFEEKFGDMHSYVRSKRLREEARLFRQ 188
 QY 166 MATALAHCHQGLVLRDLKCRFVADREKKLVLENLSDSCVLTPGDDSLMDKACPAY 225
 DB 189 IVSAVAHCHQSAIVLGDILKRFVSTERRTQRLGSLDTHIIKGEDDALSDKHGCPAY 248
 QY 226 VGPETLSRSASGKAADVWSLGVALLFTMLAGHYFODSEPVLLRGKIRGAYALPAGLS 285
 DB 249 VSPETLNTGTYSGKAADVWSLGVALLFTMLAGHYFODSEPVLLRGKIRGAYALPAGLS 308
 QY 286 APARCLVCLRLRREPARLITAGILHPMLRQDPMPLATRSHLWEAAQVVD 338
 DB 309 PPARCLIRLSLRREPSERLTAQIILHPMFEVLEP-GYVSEIGTSQDIVE 360
 RESULT 12
 Q7ZZZ7 PRELIMINARY; PRT; 344 AA.
 AC Q7ZZZ7;
 DT 01-UN-2003 (TREMBlrel. 24, Created)
 DT 01-UN-2003 (TREMBlrel. 24, last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, last annotation update)
 DE TRB-2 kinase.
 GN Xenopus laevis (African clawed frog).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Shan Y.X., Yu L.;
 RT "Cloning and characterization of the frog TRB-2 gene";
 RT Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AY254200; AAP13074.1; -.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0006474; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0004713; F:protein tyrosine kinase activity; IEA.
 DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR Pfam; PF00069; kinase; 1.
 DR PROSITE; PS00001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR SMART; SM00219; Tyr_kinase.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 DR SEQUENCE 344 AA; 38867 MW; 900F4448FPA4468E CRC64;
 Query Match 37.9%; Score 717; DB 13; Length 344;
 Best Local Similarity 46.7%; Pred. No. 4.9e-53;
 Matches 162; Conservative 45; Mismatches 122; Indels 18; Gaps 7;

DT 01-MAY-2000 (Tremblrel. 13, last sequence update)
 DT 01-JUN-2003 (Tremblrel. 24, last annotation update)
 DE Putative TRBL protein (CG5408 protein) (RH63304p).
 GN TRBL OR CG5408.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephyrididae; Drosophilidae; Drosophila.
 OK NCBI_taxid=7227;
 RN (1)
 RN SEQUENCE FROM N.A.
 RC STRAIN=Berkely;
 RC MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amaratides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale U., Bayraktoglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borzova D., Botchan M.R., Bouck U., Brokstein P., Brottier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes W., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durkin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasner K.,
 RA Glodok A., Gong F., Gorell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Katpen G.H., Ke Z., Kemnison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodly C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Switzke R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J., Q.A.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN (2)
 RN SEQUENCE FROM N.A.
 RP Seher T.C., Lepetit M.;
 RL Submitted (NCV-1999) to the EMBL/GenBank/DBJ databases.
 RN (3)
 RN SEQUENCE FROM N.A.
 RC STRAIN=y;
 RC Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Ceiniker S.;
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AEO03591; AAF51590.1; -;
 DR EMBL: AF204688; AAF26374.1; -;
 DR EMBL: BT004834; AAO45190.1; -;
 DR FlyBase: FBgn0028978; trbl.
 DR GO: GO:0000074; P:regulation of cell cycle; IMP.
 DR InterPro: IPR000719; Prot_kinase.
 DR Pfam: PF00069; kinase; 2.
 DR ProDom: PD000001; Prot_kinase; 1.

DR PROSITE: PS50011; PROTEIN KINASE_DOM; 1.
 KW Hypothetical protein; ATP-binding; Transferase.
 SQ SEQUENCE 484 AA; 5407 MW; 383BD3E5645B0D7 CRC64;
 Query Match 22.8%; Score 432; DB 5; Length 484;
 Best Local Similarity 36.1%; Freq. No. 1.6e-28;
 Matches 104; Conservative 41; Mismatches 107; Indels 36; Gaps 6;
 QY 90 TGTETCKVYVQVQAL-AVLEPYARLPKHKVAPRTEVLA-----GTOL 132
 DB 147 TGEQFLCNI--VNEPLHVCARVAFQLOQHDEELRSTVGHPLIRPVHDIPLTMDRIYI 204
 QY 133 LYA-----FFRTTGMMSLVTRRHRIPEPEAAVFRQATALAHCHGCLVRD 182
 DB 205 LIAPVQGRDSTGVTVGVYENILHTYIRHAKRLCETEARAIPHQICQVQVGRNGIILRD 264
 QY 183 LKLCRFVADDERKVLVENLEDSCVLTPGDDSLMDKACPAVVGPELSSPASYSGRKA 242
 DB 265 LKLRFFYIDARTKLQYESLEGSMTLDGEDDTSKCGCPLYTAPPELLCPQOTYKGRPA 324
 QY 243 DVMSIGVALFTMLAGHYFPDSEPVILFGKIRGAYVALPAGISAPARCIVRCLRRPEAE 302
 DB 325 DVMSIGVILYTMVQGVYFEKANCNLTIVIRHGVQVPLITSKSVRWLLSLKDKYTE 384
 QY 303 RLTATGILLHPMLRQDPMPLAPTRSHLMEAAQVVPDGLGDEAREEKG 350
 DB 385 RMTASHIFLTPLRREQ---RPFHMYLIPVDVEAED---WSDAREDEG 425

Search completed: August 24, 2004, 18:53:03
 Job time : 49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 29, 2004, 04:38:36 ; Search time 831 Seconds

(without alignments)
10817.313 Million cell updates/sec

Title: US-10-070-337-16

Perfect score: 2116

Sequence: 1 ggaagcgagcttcgcgcgcgt.....tatgaataaaaaaaaaa 2116

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseg 29Jan04:*

1: Geneseg1980s:*\n2: Geneseg1990s:*\n3: Geneseg2000s:*\n4: Geneseg2001s:*\n5: Geneseg2001bs:*\n6: Geneseg2002s:*\n7: Geneseg2003as:*\n8: Geneseg2003bs:*\n9: Geneseg2003cs:*\n10: Geneseg2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2116	100.0	2116	4 AAF30480	AAI30480 Human pro
2	2095.4	99.0	2257	9 ADCS5938	ADCS5938 DNA seque
3	2039.4	96.4	2092	4 AAI58064	AAI58064 Human pol
4	2038.6	96.3	2059	4 AAI59850	AAI59850 Human pol
5	1762.4	83.3	2048	3 AAC77866	AAC77866 Human can
6	1072.2	50.7	1077	4 AAS06709	AAS06709 Polynucle
7	1069.8	49.8	1085	3 AAZ61155	AAZ61155 cDNA J550
8	1053	45.6	1083	6 ABN86479	ABN86479 Human tri
9	803.8	38.0	972	5 AAH76218	AAH76218 Human kin
10	716.6	33.9	1076	6 ABL39747	ABL39747 Human NS
11	716.6	33.9	1076	6 ABL39762	ABL39762 Human NS
12	621.6	29.4	2004	9 ADB58995	ADB58995 Toxicity-
13	621.6	29.4	2004	9 ADB57749	ADB57749 Primary r
14	621.6	29.4	2004	9 ADCS5936	ADCS5936 DNA seque
15	562.4	25.2	996	5 AAS91231	AAS91231 DNA encod
16	533	25.2	541	6 ABR55110	ABR55110 Human col
17	505.2	23.9	2559	9 ADB31419	ADB31419 Bicalutan
18	505.2	23.9	2562	2 AAQ08817	AAQ08817 DNA encod
19	426	20.1	426	6 ABL87304	ABL87304 Human ova
20	369.4	17.5	362	6 AAC78131	AAC78131 Human can
21	354.2	16.7	396	4 AAS37461	AAS37461 Novel hum
22	315.4	14.9	353	6 ABL62004	ABL62004 Colon ade
23	297.6	14.1	534	5 AAS68600	AAS68600 DNA encod

24 269.4 12.7 3319 6 ABN86478
25 267.8 12.7 3328 9 ADD29847
26 258.4 12.2 1909 5 AAS45035
27 251.4 11.9 4221 9 ADD14683
28 249.8 11.8 3280 7 ABX63269
29 243.8 11.5 3324 6 AAS94913
30 240 11.3 690 6 AET09024
31 233.4 11.0 942 7 ABX74428
32 230.8 10.9 675 6 ABR83550
33 229.2 10.8 818 7 ACC45125
34 219 10.3 621 7 ACC45126
35 205 9.7 207 3 AAC03717
36 204.4 9.7 1943 5 AAS45223
37 165 7.8 498 6 ABR63084
38 163.6 7.7 7282 4 ABR6318
39 158 7.5 562 4 ABA61922
40 158 7.5 562 4 AAI41844
41 158 7.5 562 4 AAK36130
42 158 7.5 562 4 AAK10231
43 158 7.5 562 4 ABR35823
44 158 7.5 562 6 ABS10267
45 158 7.5 25772 4 AAK81332

ALIGNMENTS

RESULT 1
ID AAF30480 standard; cDNA; 2116 BP.
XX
AC AAF30480;
DT 29-MAY-2001 (first entry)
XX
DE Human protein phosphatase and kinase protein-5 cDNA 1271505CB1.
XX
KW Protein phosphatase and kinase protein; PPKP-5; human;
KW gastrointestinal disorder; immune system disorder; neurological disorder;
KW cell proliferative disorder; cancer; diagnosis; therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 161..1237
FT /*tag= a
FT misc_feature 326..385
FT /*tag= b
FT /note= "unique fragment"
XX
XX W0200120004-A2.
XX
XX 22-MAR-2001.
XX
XX 14-SEP-2000; 2000WC-US025515.
XX
XX 15-SEP-1999; 99US-0154141P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Yue H, Tang YT, Bandman O, Hallman JL, Baughn MR, Azimzai Y;
XX In DM;
XX WPI; 2001-244811/25.
XX DR P-PSDB; AAB20326.
XX
XX Novel human protein phosphatase and kinase proteins for diagnosis,
XX treatment and prevention of gastrointestinal, immune system, neurological
XX and cell proliferative disorders.
XX
XX Claim 5; Page 98; 103pp; English.
XX
XX The present sequence is that of cDNA encoding novel human protein

phosphatase and kinase protein PPHK-5 (see AAB20326). The cDNA was initially identified in Inyte Clone ID No. 121505C1, from a foetal lung tissue library. The PPHK-5 gene was localised to human chromosome 20. Tissues that express PPHK-5 (as a fraction of total tissues expressing PPHK-5) include reproductive (0.288), gastrointestinal (0.212) and haematopoietic or immune (0.192). Diseases or conditions associated with tissues expressing PPHK-5 (as a fraction of total tissues expressing PPHK-5) include cancer (0.577), inflammation or trauma (0.327) and cell proliferation (0.308). The encoded protein shows homology to rat kinase. The invention provides human PPHK-1 to -11 polypeptides (see AAB20322-32) and polynucleotides (see AAB20476-86). It also provides expression vectors, host cells, antibodies, agonists and antagonists, as well as methods for diagnosing, treating or preventing disorders associated with expression of PPHK, including gastrointestinal disorders, immune system disorders, neurological disorders and cell proliferative disorders, including cancer.

Sequence 2116 BP; 407 A; 663 C; 609 G; 437 T; 0 U; 0 Other;

Query Match 100.0%; Score 2116; DB 4; Length 2116;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

1  GGAGGCGGCTCCGCGCGGCTCGCTAGAGACCGGCGCAGGCTGAGCTGGGCTGGGA 60
Db  1  GGAGGCGGCTCCGCGCGGCTCGCTAGAGACCGGCGCAGGCTGAGCTGGGCTGGGA 60
Qy  1  GGAGGCGGCTCCGCGCGGCTCGCTAGAGACCGGCGCAGGCTGAGCTGGGCTGGGA 120
Db  61  TCCCGAGCTCGGCGCAGCGAGCGGCGGCGGCTGCTGCTGCTGAGCTGAGCTGAG 120
Qy  61  TCCCGAGCTCGGCGCAGCGAGCGGCGGCGGCTGCTGCTGCTGAGCTGAGCTGAG 120
Db  61  TCCCGAGCTCGGCGCAGCGAGCGGCGGCGGCTGCTGCTGCTGAGCTGAGCTGAG 120
Qy  121  CCCCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
Db  121  CCCCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
Qy  121  CCCCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
Db  121  CCCCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
Qy  181  TGCTCTGCGGCGGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db  181  TGCTCTGCGGCGGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Qy  241  CGAGGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Db  241  CGAGGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Qy  301  GTTGGCCCTGAGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db  301  GTTGGCCCTGAGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Qy  361  TCTTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db  361  TCTTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Qy  421  CTGCGCTTACAGGCACTGATGATGATGATGATGATGATGATGATGATGATGATG 480
Db  421  CTGCGCTTACAGGCACTGATGATGATGATGATGATGATGATGATGATGATGATG 480
Qy  481  GGTGAGGCGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
Db  481  GGTGAGGCGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
Qy  541  GGTGAGGCGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
Db  541  GGTGAGGCGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
Qy  601  GGTGAGGCGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
Db  601  GGTGAGGCGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
Qy  661  CACCGCGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
Db  661  CACCGCGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
Qy  721  CTTTGTCTGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780

```

```

721  CTTTGTCTGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
Qy  781  CTTTGTCTGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
Db  781  CTTTGTCTGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
Qy  841  ACCTGAGATCTGAGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
Db  841  ACCTGAGATCTGAGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
Qy  901  GGGGCTGAGCGCTTTCACATGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
Db  901  GGGGCTGAGCGCTTTCACATGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
Qy  961  CTTGCTCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
Db  961  CTTGCTCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
Qy  1021  TGCCGCTGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1080
Db  1021  TGCCGCTGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1080
Qy  1081  AGGCACTCTCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1140
Db  1081  AGGCACTCTCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1140
Qy  1141  CCATCTCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
Db  1141  CCATCTCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
Qy  1201  AGAGGAGGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
Db  1201  AGAGGAGGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
Qy  1261  TGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db  1261  TGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Qy  1321  CAAACCTTCAAGGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db  1321  CAAACCTTCAAGGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Qy  1381  ACACATCTGCTTGTTCACACACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db  1381  ACACATCTGCTTGTTCACACACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Qy  1441  CCTGTCTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Db  1441  CCTGTCTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Qy  1501  GAGATGACAACTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 1560
Db  1501  GAGATGACAACTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 1560
Qy  1561  TACACTGAGTACACTTGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
Db  1561  TACACTGAGTACACTTGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
Qy  1621  TGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
Db  1621  TGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
Qy  1681  GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Db  1681  GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Qy  1741  AAGCCAGCTCACTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
Db  1741  AAGCCAGCTCACTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
Qy  1801  CTCCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
Db  1801  CTCCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860

```

QY 1861 GTTCTGTCMACTCAAGACTGTTCTGGAATGAGGGTCCAGGCTGTCAACCATGGGGCT 1920
 Db 1861 GTTCTGTCMACTCAAGACTGTTCTGGAATGAGGGTCCAGGCTGTCAACCATGGGGCT 1920
 QY 1921 TCTGACCTGAGCACCAGGTTGAGGAGATTAGGAGGGTCTGCTGTGGCACCCT 1980
 Db 1921 TCTGACCTGAGCACCAGGTTGAGGAGATTAGGAGGGTCTGCTGTGGCACCCT 1980
 QY 1981 GGAAGTCCGAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACATCCAGGTCCATAC 2040
 Db 1981 GGAAGTCCGAGGTGGGACTCTTCTGGGGACACTTGGGGTCCACATCCAGGTCCATAC 2040
 QY 2041 TCTAGTTTGGATCCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2100
 Db 2041 TCTAGTTTGGATCCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2100
 QY 2101 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2116
 Db 2101 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2116
 RESULT 2
 ID ADC59338
 AC ADC59338 standard; DNA; 2257 BP.
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE DNA sequence encoding human cell death inhibitory protein.
 XX
 KW Human; cell-death; cardiact; neuroprotective; anti-HIV; antiinflammatory;
 KW cerebroprotective; hepatotropic; cytosolic; immunosuppressive;
 KW antirheumatic; ophthalmological; nootropic; antiparkinsonian;
 KW anticonvulsant; hypotensive; antiarteriosclerotic; haemostatic;
 KW antialcoholic; virucide; HIV; cardiac disease; immunological disease;
 KW neurodegenerative disease; ischaemic damage;
 KW central nervous system disorder; hepatic disorder;
 KW acute pancreatitis; inflammation; cancer; AIDS; autoimmune disease;
 KW rheumatoid; Crohn's disease; glaucoma; Alzheimer's disease;
 KW Parkinson's disease; Huntington's disease; hypertension;
 KW arteriosclerosis; reperfusion damage; myocardial infarction;
 KW cerebral trauma; cerebral infarction; cerebral haemorrhage; hepatitis;
 KW alcoholic hepatitis; cerebral ischemia; ss; gene.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 295..1371
 FT /*tag= a
 FT /product= "Human cell death inhibitory protein"
 XX
 PN JP2003063986-A.
 XX
 PD 05-MAR-2003.
 XX
 PF 27-AUG-2001; 2001JP-00255811.
 XX
 PR 27-AUG-2001; 2001JP-00255811.
 XX
 PA (SUMU) SUMITOMO SEIYAKU KK.
 XX
 DR WPI; 2003-601360/57.
 DR P-PSDB; ADC59337.
 XX
 PT Controlling cell-death by administering positive or negative regulator of
 PT cell-death inhibitory-factor, for treating cancer; AIDS; autoimmune
 PT diseases, Crohn's diseases, glaucoma, Alzheimer disease.
 PS Claim 1; SEQ ID NO 4; 35pp; Japanese.
 XX
 CC This invention relates to a novel method for controlling cell-death

CC comprising administering a positive or negative regulator of cell-death
 CC inhibitory-factor. The invention also comprises a method for screening
 CC modulators of cell-death, by contacting cells expressing cell-death
 CC inhibitory factor with a candidate compound, monitoring level of
 CC expression of cell-death inhibitory factor, evaluating cell-death
 CC modulation ability of the compound based on change in the level of
 CC expression of the factor and selecting compounds having cell-death
 CC modulation ability. The cell death regulator of the invention may have
 CC cardiact, neuroprotective, anti-HIV, antiinflammatory, cerebroprotective,
 CC hepatotropic, cytosolic, immunosuppressive, antirheumatic,
 CC ophthalmological, nootropic, antiparkinsonian, anticonvulsant,
 CC hypotensive, antiarteriosclerotic, haemostatic, antialcoholic and
 CC virucide activities. The method of the invention is useful for treating
 CC HIV, cardiac diseases, immunological diseases, neurodegenerative disease,
 CC ischaemic damage and congestion, disorder of central nervous system,
 CC hepatic disorder, acute pancreatitis, inflammation, and cancer, AIDS,
 CC autoimmune diseases, rheumatism, Crohn's diseases, glaucoma, Alzheimer's
 CC disease, Parkinson's disease, Huntington's disease, hypertension,
 CC arteriosclerosis, reperfusion damage, myocardial infarction, cerebral
 CC trauma, cerebral infarction, cerebral haemorrhage, hepatitis, alcoholic
 CC hepatitis, and cerebral ischemia. The present sequence represents the DNA
 CC sequence encoding the human cell death inhibitory protein of the
 CC invention.
 XX
 SQ Sequence 2257 BP; 435 A; 705 C; 659 G; 458 T; 0 U; 0 Other;

Query Match 99.0%; Score 2095.4; DB 9; Length 2257;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 2110; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 1 GAGAGCGGCTCCGGCGCGGCTGCTGAGAACCGGGCAGGGCTGAGCTGGGCTGGGA 60
 Db 134 GAGAGCGGCTCCGGCGCGGCTGCTGAGAACCGGGCAGGGCTGAGCTGGGCTGGGA 193
 QY 61 TCCGAGGTGGGAG 119
 Db 194 TCCGAGGTGGGAG 253
 QY 120 GCCCGGCGGCG 179
 Db 254 GCCCGGCGGCG 313
 QY 180 CTGCTCTGCGGAGTTCCTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 239
 Db 314 CTGCTCTGCGGAGTTCCTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 373
 QY 240 CCGAGCGTCCCGTCCAG 299
 Db 374 CCGAGCGTCCCGTCCAG 433
 QY 300 TGTGCGCGTGGAGCGCACTACTGCTCCAGATGCGCAACTGCTGTGGCGCACTGCTGCC 359
 Db 434 TGTGCGCGTGGAGCGCACTACTGCTCCAGATGCGCAACTGCTGTGGCGCACTGCTGCC 493
 QY 360 GTCTTGGGCGCTATGTCCTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 419
 Db 494 GTCTTGGGCGCTATGTCCTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 553
 QY 420 ACTGCGCTACAGGAG 479
 Db 554 ACTGCGCTACAGGAG 613
 QY 480 TGTGAGAGCGCTACAGGAG 539
 Db 614 TGTGAGAGCGCTACAGGAG 673
 QY 540 TGGCTGGTACCGAGCTCTCTACAGCGCTTTTTCATCTGAGAGAGAGAGAGAGAGAGAG 599
 Db 674 TGGCTGGTACCGAGCTCTCTACAGCGCTTTTTCATCTGAGAGAGAGAGAGAGAGAGAG 733
 QY 600 TGGTGGAGAGCGGAG 659
 Db 734 TGGTGGAGAGCGGAG 793

QY 660 CCACCGCCCTGGCGCATCTGTCAACGACAGCTTGTCTGTCTGTATCTCAAGCTGTGTCTC 719
 Db 794 CCACCGCCCTGGCGCATCTGTCAACGACAGCTTGTCTGTCTGTATCTCAAGCTGTGTCTC 853
 QY 720 GCTTGTCTGTCTGTCAACGACAGCTTGTCTGTCTGTATCTCAAGCTGTGTCTC 779
 Db 854 GCTTGTCTGTCTGTCAACGACAGCTTGTCTGTCTGTATCTCAAGCTGTGTCTC 913
 QY 780 GCGTGTCTGTCTGTCAACGACAGCTTGTCTGTCTGTATCTCAAGCTGTGTCTC 839
 Db 914 GCGTGTCTGTCTGTCAACGACAGCTTGTCTGTCTGTATCTCAAGCTGTGTCTC 973
 QY 840 GACCTGAGTACTCTGAGCTGAGGAGCTCTCACTCTGAGGACAGGACGCTTGTCTGAGCC 899
 Db 974 GACCTGAGTACTCTGAGCTGAGGAGCTCTCACTCTGAGGACAGGACGCTTGTCTGAGCC 1033
 QY 900 TGGGCGTGGCGCTCTTCAACATGCTGAGCGGACCTAACCCTTCCAGAGCTGAGCCCTG 959
 Db 1034 TGGGCGTGGCGCTCTTCAACATGCTGAGCGGACCTAACCCTTCCAGAGCTGAGCCCTG 1093
 QY 960 TCTGTCTCTTGGGCAAGATCCGCGGCGGCTTACGCTTGTCTGTCTGAGGCTCTGTGCCC 1019
 Db 1094 TCTGTCTCTTGGGCAAGATCCGCGGCGGCTTACGCTTGTCTGTCTGAGGCTCTGTGCCC 1153
 QY 1020 CTGGCCGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 1079
 Db 1154 CTGGCCGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 1213
 QY 1080 CAGGCAATCTCTCTGACCCCTGGCTGCGACAGGACCCGATGCTTGTGAGCCGAT 1139
 Db 1214 CAGGCAATCTCTCTGACCCCTGGCTGCGACAGGACCCGATGCTTGTGAGCCGAT 1273
 QY 1140 CCCATCTCTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1199
 Db 1274 CCCATCTCTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1333
 QY 1200 AAGAGAGAGAGACAG 1259
 Db 1334 AAGAGAGAGAGACAG 1393
 QY 1260 CTGCAACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1319
 Db 1394 CTGCAACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1453
 QY 1320 CCAAACTTTCAGTCTTCCAG 1379
 Db 1454 CCAAACTTTCAGTCTTCCAG 1513
 QY 1380 TACACATCTGCTTGTTCAC 1439
 Db 1514 TACACATCTGCTTGTTCAC 1573
 QY 1440 GCGCTGTCTCGGTGTGAG 1499
 Db 1574 GCGCTGTCTCGGTGTGAG 1633
 QY 1500 AGAGATGACAACTGGCATCTTGAAGCTGACACACCTTTCATGACATAGAGTCACTGT 1559
 Db 1634 AGAGATGACAACTGGCATCTTGAAGCTGACACACCTTTCATGACATAGAGTCACTGT 1693
 QY 1560 CTACACCTGGGTACCTTGTGACAGAGTGGGCTCCACATGATCTGAGGAGAGAGAGAGAG 1619
 Db 1694 CTACACCTGGGTACCTTGTGACAGAGTGGGCTCCACATGATCTGAGGAGAGAGAGAGAG 1753
 QY 1620 CTGTCCAGAGACATCCCTTTCACAAACAAACAGCTGCTTGTATCTTGTATCTTGTATCTT 1679
 Db 1754 CTGTCCAGAGACATCCCTTTCACAAACAAACAGCTGCTTGTATCTTGTATCTTGTATCTT 1813
 QY 1680 AGAGAAAGGAGATCTCTGTGCAAGAGCTCCAGGCTCTCTCTCTGCACTCAGAGACC 1739
 Db 1814 AGAGAAAGGAGATCTCTGTGCAAGAGCTCCAGGCTCTCTCTCTGCACTCAGAGACC 1873

QY 1740 CAAGCCAGCTCACTCTGGAGACGTGTCTCCAGACATCTGTCTCTGTATTAAGAGAT 1799
 Db 1874 CAAGCCAGCTCACTCTGGAGACGTGTCTCCAGACATCTGTCTCTGTATTAAGAGAT 1933
 QY 1800 TCTCTTCCAGGCTTAAGCTTGGGATTTGGGCGAGAGATTAAGATCCAACTATGAGGCT 1859
 Db 1934 TCTCTTCCAGGCTTAAGCTTGGGATTTGGGCGAGAGATTAAGATCCAACTATGAGGCT 1993
 QY 1860 AGTCTTGTCTTAACCTCAAGACCTGTCTGGAATGAGGGTCCAGGCTGTCAACATGAGGCT 1919
 Db 1994 AGTCTTGTCTTAACCTCAAGACCTGTCTGGAATGAGGGTCCAGGCTGTCAACATGAGGCT 2053
 QY 1920 TTCTGACCTGAGCACCAAGATTTGAGAGACAGATTAAGAGAGGCTGTCTGTGACACC 1979
 Db 2054 TTCTGACCTGAGCACCAAGATTTGAGAGACAGATTAAGAGAGGCTGTCTGTGACACC 2113
 QY 1980 TGGAAAGTCCAGGAGGAGCTTTCTGGGAGACATTTGGGCTCAAAATCCAGGCTCCANA 2039
 Db 2114 TGGAAAGTCCAGGAGGAGCTTTCTGGGAGACATTTGGGCTCAAAATCCAGGCTCCANA 2173
 QY 2040 CTCTAGGTTTGGATACCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2099
 Db 2174 CTCTAGGTTTGGATACCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2233
 QY 2100 GAAATAAAAA 2116
 Db 2234 GAAATAAAAA 2250

RESULT 3
 ID AAI58064
 AAI58064 standard; cDNA; 2092 BP.

AAI58064;

22-OCT-2001 (first entry)

Human polynucleotide SEQ ID NO 267.

Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 peripheral nervous system; neuropathy; central nervous system; CNS;
 Alzheimer's; Parkinson's disease; Huntington's disease; hemostatic;
 amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 leukemia; ss.

Homo sapiens.

WO200153312-A1.

26-JUL-2001.

26-DEC-2000; 2000WO-US034263.

23-DEC-1999; 99US-00471275.

21-JAN-2000; 2000US-00486725.

25-APR-2000; 2000US-00523117.

20-JUN-2000; 2000US-00598042.

19-JUL-2000; 2000US-00620312.

03-AUG-2000; 2000US-00653450.

14-SEP-2000; 2000US-00662191.

19-OCT-2000; 2000US-00693036.

29-NOV-2000; 2000US-00727344.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D, Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao Q, Zhou P, Goodrich R, Dmanac R, WPI, 2001-442253/47. P-PSDB; AAM38908.

QY 1838 TAAGATCAAACTATGAGGCTAGTCTTGTCTAACTCAAGCTGTTCTGAAATGAGGCT 1897
DB 1816 TAAGATCAAACTATGAGGCTAGTCTTGTCTAACTCAAGCTGTTCTGAAATGAGGCT 1875
QY 1898 CAGAGCTGTAAACATGAGGCTTCTGACCTGAGACCAAGGTTGAGGAGACAGATTAGG 1957
DB 1876 CAGAGCTGTAAACATGAGGCTTCTGACCTGAGACCAAGGTTGAGGAGACAGATTAGG 1935
QY 1958 CAGAGCTGTCTGTGAGGCTGACCTGAGAAAGTCCAGAGTGGAGCTCTTCTGAGGACACTTGG 2017
DB 1936 CAGAGCTGTCTGTGAGGCTGACCTGAGAAAGTCCAGAGTGGAGCTCTTCTGAGGAGCACTTGG 1995
QY 2018 GGTCCCAATCCCAAGCTGCAATCTCTAGGTTTGGATACCATGATGATGTTTACTT 2077
DB 1996 GGTCCCAATCCCAAGCTGCAATCTCTAGGTTTGGATACCATGATGATGTTTACTT 2055
QY 2078 GTGCTTAATAAGGAGATTAATAATAAAAAAAAAA 2114
DB 2056 GTGCTTAATAAGGAGATTAATAATAAAAAAAAAA 2092

RESULT 4
AA159850
ID AA159850 standard; cDNA; 2059 BP.
AC AA159850;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 3839.
XX
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO20015312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-0055317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
XX (HYSE-) HYSEQ INC.
XX
PA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QH;
PI Zhou P, Goodrich R, Drmanac RT;
XX
XX MPI; 2001-442253/47.
XX F-PDB; AAM40694.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders such
XX as central nervous system injuries.
XX
XX Claim 1; SEQ ID NO 3839; 10078bp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and the
XX encoded polypeptides (AAM38642-AAM42213) with nootropic,

CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 2059 BP; 400 A; 645 C; 583 G; 431 T; 0 U; 0 Other;
XX
Query Match 96.3%; Score 2038.6; DB 4; Length 2059;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2052; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 59 GATCCCGAGCTGCGAGAGAGGCA-CGGGCGGGCCACCTGCTGGTCCCTGAGAGCTCT 117
DB 3 GATCCCGAGCTGCGAGAGAGGCA-CGGGCGGGCCACCTGCTGGTCCCTGAGAGCTCT 62
QY 118 GAGCCCCGCGCGCCCGCGCCCAAGCGGAGCAAGCGGGCGAGATGAGCCACCCCTCT 177
DB 63 GAGCCCCGCGCGCCCGCGCCCAAGCGGAGCAAGCGGGCGAGATGAGCCACCCCTCT 122
QY 178 GGCTGCTCTGCGGGGTTCCCTGCTGCGAGAAAGAGGTTGAGATTGATGAACTTAGA 237
DB 123 GGCTGCTCTGCGGGGTTCCCTGCTGCGAGAAAGAGGTTGAGATTGATGAACTTAGA 182
QY 238 TACCGAGGTCCTCCGTCGAGAAACGAGTCGAAAGTGGGCCCCAGCCAGCTGCCCCCTG 297
DB 183 TACCGAGGTCCTCCGTCGAGAAACGAGTCGAAAGTGGGCCCCAGCCAGCTGCCCCCTG 242
QY 298 CCTGTTGCCCTGAGCCCACTACTGCTCCAGATGTCGCAACTGCTGAGCCACTGCTC 357
DB 243 CCTGTTGCCCTGAGCCCACTACTGCTCCAGATGTCGCAACTGCTGAGCCACTGCTC 302
QY 358 CCGTCTTGGGCCCTATATGCTCTCTGAGAGCCGAGAGAGGCGGGGCTTACCAAGGCCCT 417
DB 303 CCGTCTTGGGCCCTATATGCTCTCTGAGAGCCGAGAGAGGCGGGGCTTACCAAGGCCCT 362
QY 418 GCATGCTCCTTAAGCACTGATATACCTGCAAGGTTGATCCCGTCAGAGAGCCCTGAC 477
DB 363 GCATGCTCCTTAAGCACTGATATACCTGCAAGGTTGATCCCGTCAGAGAGCCCTGAC 422
QY 478 CGTGTGAGAGCCCTAAGCGAGGCTGCCCGCGCAAGCATGTGGCTGGCCCACTGAGGT 537
DB 423 CGTGTGAGAGCCCTAAGCGAGGCTGCCCGCGCAAGCATGTGGCTGGCCCACTGAGGT 482
QY 538 CCTGAGTGTATCCAGAGCTCTCTAAGCCTTTTCACTCGGACCCATGGGGAATGTGACG 597
DB 483 CCTGAGTGTATCCAGAGCTCTCTAAGCCTTTTCACTCGGACCCATGGGGAATGTGACG 542
QY 598 CCTGAGTGTATCCAGAGCTCTCTAAGCCTTTTCACTCGGACCCATGGGGAATGTGACG 657
DB 543 CCTGAGTGTATCCAGAGCTCTCTAAGCCTTTTCACTCGGACCCATGGGGAATGTGACG 602
QY 658 GAGCAACGCGCTGAGCGCACTGTCAACGAGACGAGTGTCTCGGTGATCTCAAGCTGTG 717
DB 603 GAGCAACGCGCTGAGCGCACTGTCAACGAGACGAGTGTCTCGGTGATCTCAAGCTGTG 662
QY 718 TCGCTTGTCTTGTGCTGACCCGTGAGAGAAAGCTGTGTGTGAGAAACCTGAGAGACTC 777
DB 663 TCGCTTGTCTTGTGCTGACCCGTGAGAGAAAGCTGTGTGTGAGAAACCTGAGAGACTC 722
QY 778 CTGCGGTGAGCACTGGGCGAGATATCTCTGTTGGACAAGACGCGCTGCGACCTTACGT 837
DB 723 CTGCGGTGAGCACTGGGCGAGATATCTCTGTTGGACAAGACGCGCTGCGACCTTACGT 782
QY 838 GAGACCTGAGATCTAGCTCAGGGGCTCATATCTGCGGCAAGGACCGAGATCTTGAG 897

Db 783 GGGACCTGAGATGCTCAGGCTCAGGGGCTATCTGGGGAAGGACCGATGCTGGAG 842
 QY 898 CCTGGGCGTGGCGCTCTTACCAATGCTGCGGCGCACTAACCTTCCAGAGCTCGAGCC 957
 Db 843 CCGGGGCGTGGCGCTCTTACCAATGCTGCGGCGCACTAACCTTCCAGAGCTCGAGCC 902
 QY 958 TGTGCTGCTTGGGGAAGATCGCGCGGCGGCTTACGCTTGGCTGCAAGGCTCTCGGC 1017
 Db 903 TGTGCTGCTTGGGGAAGATCGCGCGGCGGCTTACGCTTGGCTGCAAGGCTCTCGGC 962
 QY 1018 CCTGCGCGCTGTCTGGTTGCTGCTGCTCTGCTGCTGCGGAGCGAGTGAAGGCTCAGC 1077
 Db 963 CCTGCGCGCTGTCTGGTTGCTGCTGCTCTGCTGCTGCGGAGCGAGTGAAGGCTCAGC 1022
 QY 1078 CACAGGCAATCTCTGCGACCCCTGCTGCGACAGGACCCGATGCTTACCCCAACCG 1137
 Db 1023 CACAGGCAATCTCTGCGACCCCTGCTGCGACAGGACCCGATGCTTACCCCAACCG 1082
 QY 1138 ATCCCAATCTCTGCGAGGCTGCGGCGGCTGCTGCTGCTGCTGCGAGGCTGCGAG 1197
 Db 1083 ATCCCAATCTCTGCGAGGCTGCGGCGGCTGCTGCTGCTGCTGCGAGGCTGCGAG 1142
 QY 1198 GGAAGAGGAGGAGACAGAGAGTGGTTCTGTATGCTAGGACCAACCCCTACTACAGCTC 1257
 Db 1143 GGAAGAGGAGGAGACAGAGAGTGGTTCTGTATGCTAGGACCAACCCCTACTACAGCTC 1202
 QY 1258 AGCTGCAACAGTGAATGATTTGGGGGTAGCTCCAGGCTTCTCTGCTGCTGCAATG 1317
 Db 1203 AGCTGCAACAGTGAATGATTTGGGGGTAGCTCCAGGCTTCTCTGCTGCTGCAATG 1262
 QY 1318 AACCAAACTCTAGTGGCTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1377
 Db 1263 AACCAAACTCTAGTGGCTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1322
 QY 1378 TGTACACATCTGCTTGTGTCCACACATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1437
 Db 1323 TGTACACATCTGCTTGTGTCCACACATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1382
 QY 1438 AAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAAGCAAAAGAGCAAAATTTCCCTGCTC 1497
 Db 1383 AAGCCCTGTTCTCGGTGCTGGGAGTACAGCAGTGAAGCAAAAGAGCAAAATTTCCCTGCTC 1442
 QY 1498 ACAGAGATGACAACTGGAGATCCTGAGTGAACAACCTTTCATGACCATAGTCACT 1557
 Db 1443 ACAGAGATGACAACTGGAGATCCTGAGTGAACAACCTTTCATGACCATAGTCACT 1502
 QY 1558 GTCTACCTGGGTACACTTGTGTACCAAGATGTCGCTCCACATGATGCTGCTCAGGAC 1617
 Db 1503 GTCTACCTGGGTACACTTGTGTACCAAGATGTCGCTCCACATGATGCTGCTCAGGAC 1562
 QY 1618 CTCTGTCCAAAGCAATCCCTTTCACAAACCAAGCTGCTTGTGTATCTGTACCTTT 1677
 Db 1563 CTCTGTCCAAAGCAATCCCTTTCACAAACCAAGCTGCTTGTGTATCTGTACCTTT 1622
 QY 1678 TCAGAGAAAGGAGGATCCCTGTGTCCAAAGGCTCAGGCTCTCTCCCTGCACTCAGGA 1737
 Db 1623 TCAGAGAAAGGAGGATCCCTGTGTCCAAAGGCTCAGGCTCTCTCCCTGCACTCAGGA 1682
 QY 1738 CCCAAGCCAGCTCACTCTGGGAATGTTGCCAGCACTCTGCTCTGTGATTAAG 1797
 Db 1683 CCCAAGCCAGCTCACTCTGGGAATGTTGCCAGCACTCTGCTCTGTGATTAAG 1742
 QY 1798 ATTCTCTTCCAGGCTTAAGCTGGGATTTGGGCCAGAGATTAAGATCAAACTATGAG 1857
 Db 1743 ATTCTCTTCCAGGCTTAAGCTGGGATTTGGGCCAGAGATTAAGATCAAACTATGAG 1802
 QY 1858 CTAGTCTTGTCTAATCTAAGATCTTGTGAATGAGGCTCAGGCTCTGTCAACATGG 1917
 Db 1803 CTAGTCTTGTCTAATCTAAGATCTTGTGAATGAGGCTCAGGCTCTGTCAACATGG 1862
 QY 1918 GCTTCTGACTGAGCACCAGGTTGAGGAGCAGATTTAGGAGGAGGTTGCTCTGTGGCCA 1977

Db 1863 GCTTCTGACCTGAGACCAAGTTGAGGAGCAGGATTAAGCAGGCTGTCTGTGGCCA 1922
 QY 1978 CCTGGAAGTCCAGGTGAGACTTCTGGGAGACATTTGGGGTCCAAATCCAGGTCCA 2037
 Db 1923 CCTGGAAGTCCAGGTGAGACTTCTGGGAGACATTTGGGGTCCAAATCCAGGTCCA 1982
 QY 2038 TACTCTAGGTTTGGATACCATGATGATGATGATGATGATGATGATGATGATGAT 2097
 Db 1983 TACTCTAGGTTTGGATACCATGATGATGATGATGATGATGATGATGATGATGAT 2042
 QY 2098 ATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2114
 Db 2043 ATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2059
 RESULT 5
 AAC77866
 ID AAC77866 standard; cDNA; 2048 BP.
 XX
 AC AAC77866;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human cancer associated gene sequence SEQ ID NO:260.
 KW Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnaric; immunomodulatory;
 KW antidiabetic; antiaesthetic; antirheumatic; antiarthritic; antiviral;
 KW antiinflammatory; antihypertoid; antiallergic; antibacterial; cardiant;
 KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening; ss.
 OS Homo sapiens.
 PN WO20055350-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US005862.
 XX
 PR 12-MAR-1999; 99US-0124270P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI; 2000-587533/55.
 XX
 DR P-PSDB; AAB43657.
 PT Novel isolated nucleic acids comprising sequences encoding peptides
 useful for treating or diagnosing e.g. cancer.
 PS Claim 1; Page 825-826; 2352zp; English.
 CC AAC77866 to AAC77848 encode the human cancer associated proteins given in
 CC AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnaric; immunomodulator;
 CC antidiabetic; antiaesthetic; antirheumatic; antiarthritic;
 CC antiinflammatory; antihypertoid; antiallergic; antibacterial; antiviral;
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
 CC nootropic; vasotropic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ

RESULT 6
 AAS06709
 ID AAS06709 standard; cDNA; 1077 BP.
 XX
 AC AAS06709;
 XX
 DT 12-SEP-2001 (first entry)
 XX
 DE Polynucleotide sequence encoding human protein kinase #9.
 XX
 KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
 KW metabolic disorder; immune related disease; neurological disorder;
 KW neurodegenerative disorder; inflammatory disorder; infectious disease;
 KW reproductive disorder; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200138503-A2.
 XX
 PD 31-MAY-2001.
 XX
 PF 22-NOV-2000; 2000WO-US032085.
 XX
 PR 24-NOV-1999; 99US-0167482P.
 XX
 PA (SUGEN-) SUGEN INC.
 XX
 PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
 PI Flanagan P, Clary D;
 XX
 DR WPI; 2001-343950/36.
 XX
 P-PSDB; AAU03509.
 XX
 PT Nucleic acids encoding human kinase polypeptides, useful for preventing
 PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
 PT neuronal-associated diseases, and microbial infections.
 XX
 PS Example 1; Fig 1; 433bp; English.
 XX
 CC AAS06701-AAS06757 encode for novel human protein kinases #1-57. The novel
 CC protein kinases have been identified as members of the tyrosine or
 CC serine/threonine kinase (PTK and STK) families. The polynucleotides
 CC encoding protein kinases and the polypeptides may be used in the
 CC prevention, diagnosis and treatment of diseases associated with
 CC inappropriate kinase expression. For example, they may be used to treat
 CC cancers (especially cancers of haematopoietic origin), cardiovascular
 CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
 CC immune related diseases (e.g. rheumatoid arthritis), neurological
 CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
 CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
 CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
 CC Additionally, polynucleotides encoding protein kinases may be used for
 CC gene therapy and as DNA probes in diagnostic assays. The protein kinase
 CC polypeptides may be used as antigens in the production of antibodies
 CC against the protein kinases and in assays to identify modulators of
 CC protein kinase expression and activity
 XX
 SQ Sequence 1077 BP; 174 A; 373 C; 325 G; 205 T; 0 U; 0 Other;
 XX

Query Match 50.7%; Score 1072.2; DB 4; Length 1077;
 Best Local Similarity 99.7%; Pred. No. 2e-259;
 Matches 1074; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 161 ATGGAGCCACCCCTCTGGCTGCTCTCGGCGTTCCCTGTCCAGAAAGCGGTTGGAG 220
 DB 1 ATGGAGCCACCCCTCTGGCTGCTCTCGGCGTTCCCTGTCCAGAAAGCGGTTGGAG 60
 QY 221 TTGGATGACAACTAGATACCGAGCGTCCCTCCAGAAAGAGCTCCAGAGTGGGCCCCAG 280
 DB 61 TTGGATGACAACTAGATACCGAGCGTCCCTCCAGAAAGAGCTCCAGAGTGGGCCCCAG 120
 QY 281 CCCAGACTGCCCCCTGCTGTGCTCCCTGAGCCCACTACTGCTCCAGATCGTGCAGCT 340
 |||||||

DB 121 CCCAGACTGCCCCCTGCTGTGCCCCCTGAGCCCACTACTGCTCCAGATCGTGCAGCT 180
 QY 341 GCTGTGGCCACTGACCTCCCGTCTGTGGGCCCTTATGTCCTCTGAGGCCAGAGAGGGCCGG 400
 DB 181 GCTGTGGCCACTGACCTCCCGTCTGTGGGCCCTTATGTCCTCTGAGGCCAGAGAGGGCCGG 240
 QY 401 CGGGCCCTACCAAGCCCTGCACTGACCTTACAGGCACTGATATACCTGCAAGGTATACCC 460
 DB 241 CGGGCCCTACCAAGCCCTGCACTGACCTTACAGGCACTGATATACCTGCAAGGTATACCC 300
 QY 461 GTCCAGAAAGCCCTGCGCTGCTGAGACCTTACCGCGGCTGCCGCCCAAGACATGTG 520
 DB 301 GTCCAGAAAGCCCTGCGCTGCTGAGACCTTACCGCGGCTGCCGCCCAAGACATGTG 360
 QY 521 GCTCGGCCCACTGAGGTCCTGGCTGCTGATCCGACCTCTTACGCTTTTCACTCGGACC 580
 DB 361 GCTCGGCCCACTGAGGTCCTGGCTGCTGATCCGACCTCTTACGCTTTTCACTCGGACC 420
 QY 581 CATGGGGAACATGACAGCCTGTGCGAAACGCGCCACCGTATCCCTGAGCCTGAGGCTGCG 640
 DB 421 CATGGGGAACATGACAGCCTGTGCGAAACGCGCCACCGTATCCCTGAGCCTGAGGCTGCG 480
 QY 641 GTGCTCTTCCGCAATGAGCCACCGCCTGTGCGCACTGTACCAAGACGCTGTGCTCG 700
 DB 481 GTGCTCTTCCGCAATGAGCCACCGCCTGTGCGCACTGTACCAAGACGCTGTGCTCG 540
 QY 701 CGTGAATCAAGCTGTGTGCTTGTCTTGTGCTGACCGGTGAGAGAAAGGTGTGCTCG 760
 DB 541 CGTGAATCAAGCTGTGTGCTTGTCTTGTGCTGACCGGTGAGAGAAAGGTGTGCTCG 600
 QY 761 GAGAACTGAGAGACTCTGTGCTGCTGACTGAGGCGCAGATGATTCCTGTGAGCAAGAC 820
 DB 601 GAGAACTGAGAGACTCTGTGCTGCTGACTGAGGCGCAGATGATTCCTGTGAGCAAGAC 660
 QY 821 GCGTCCCAAGCCTTACGTGAGGACCTGAGATCTCACTCAAGGCTCTATCTCGGCAAG 880
 DB 661 GCGTCCCAAGCCTTACGTGAGGACCTGAGATCTCACTCAAGGCTCTATCTCGGCAAG 720
 QY 881 GCAGCCGATGCTGAGAGCTGTGGGCTGTGGCGCTTCAACATGCTGTGGCCCACTACCC 940
 DB 721 GCAGCCGATGCTGAGAGCTGTGGGCTGTGGCGCTTCAACATGCTGTGGCCCACTACCC 780
 QY 941 TTCCAGACCTGAGAGCTGTCTGTCTTCCGCAAGATCCGCGGGGCTTACGCTTG 1000
 DB 781 TTCCAGACCTGAGAGCTGTCTGTCTTCCGCAAGATCCGCGGGGCTTACGCTTG 840
 QY 1001 CCTGAGGCTCTGTGGCCCTGCGCGCTGTCTGTGCTGCTCTTCTGTGAGAGCA 1060
 DB 841 CCTGAGGCTCTGTGGCCCTGCGCGCTGTCTGTGCTGCTCTTCTGTGAGAGCA 900
 QY 1061 GCTGAACGGCTCAAGCAAGCAGCATCTCTGACCCCTGGCTGGCAAGAGCCGATG 1120
 DB 901 GCTGAACGGCTCAAGCAAGCAGCATCTCTGACCCCTGGCTGGCAAGAGCCGATG 960
 QY 1121 CCTTTAGCCCAACCCGATCCATCTCTGAGGCTGCCAGAGTGTCTCCATGATGACTG 1180
 DB 961 CCTTTAGCCCAACCCGATCCATCTCTGAGGCTGCCAGAGTGTCTCCATGATGACTG 1020
 QY 1181 GGGCTGAGCAAGCCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1237
 DB 1021 GGGCTGAGCAAGCCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1077

RESULT 7
 AAZ61155
 ID AAZ61155 standard; cDNA; 1085 BP.
 XX
 AC AAZ61155;
 XX
 DT 30-MAY-2000 (first entry)
 XX
 DE cDNA J0503-KS encoding domains VIA to XI of a protein kinase.
 XX


```

XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT CDS 1..972
XX FT /tag= a
XX FT /product= "PKIN-10"
XX PN MO200160991-A2.
XX PD 23-AUG-2001.
XX PF 16-FEB-2001; 2001WO-US005240.
XX PR 17-FEB-2000; 2000US-0183682P.
XX PR 02-MAR-2000; 2000US-0186559P.
XX PR 09-MAR-2000; 2000US-018606P.
XX PR 17-MAR-2000; 2000US-0189989P.
XX PR 30-MAR-2000; 2000US-0193851P.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Tang YT, Buford N, Gandhi AR, Patterson C, Khan FA, Yue H;
XX PI Hafalia A, Shih IL, Tribouley CM, Yao MG, Buttrill JD, Marcus GA;
XX PI Zingler KA, Lu DM, Bandman O, Policky JL, Griffin JA, Thornton M;
XX PI Nguyen DB, Lal P, Walsh RT;
XX DR WPI; 2001-514771/56.
XX DR P-PSDB; AAB85791.
XX PT Isolated human kinase polypeptides useful in the diagnosis, treatment and
XX PT prevention of cancer, immune disorders and disorders affecting growth and
XX PT development.
XX PS Claim 5; Page 125; 126pp; English.
XX CC The invention provides human kinases (PKIN) and polynucleotides encoding
XX CC PKIN. The PKIN polypeptides can be expressed using standard recombinant
XX CC methodology. The PKIN polypeptides, polynucleotides, modulators and
XX CC specific antibodies are useful in the diagnosis, treatment and prevention
XX CC of cancer, immune disorders, disorders affecting growth and development,
XX CC atherosclerosis, and other cardiovascular diseases, and lipid disorders
XX CC and in the assessment of the effects of exogenous compounds on the
XX CC expression of nucleic acid sequences of human kinases. The present
XX CC sequence represents a cDNA encoding a human PKIN-10 polypeptide
XX SQ Sequence 972 BP; 167 A; 329 C; 294 G; 182 T; 0 U; 0 Other;
XX
Query Match 38.0%; Score 803.8; DB 5; Length 972;
Best Local Similarity 87.7%; Pred. No. 6,7e-192;
Matches 945; Conservative 0; Mismatches 27; Indels 105; Gaps 2;
XX
QY 161 ATGCGAGCCACCCCTCTGCTGCTGCGGGTTCCCTGTCCAGAAAGCGGTTGGAG 220
DB 1 ATGCGAGCCACCCCTTTGGCTGCTGCGGGTTCCCTGTCCAGAAAGCGGTTGGAG 60
QY 221 TTGGATGACAACTTAAGTACGAGCGTCCGTCAGAAAGAGCTGGAATGGGCCCCAG 280
DB 61 TTGGATGACAACTTAAGTACGAGCGTCCGTCAGAAAGAGCTGGAATGGGCCCCAG 120
QY 281 CCCGAGCTGCCCCCTGCTGCTGTCGCGGAGCCCACTACTCTCCAGATCGTCAACT 340
DB 121 CCCGAGCTGCCCCCTGCTGCTGTCGCGGAGCCCACTACTCTCCAGATCGTCAACT 180
QY 341 GCTGTGGCACTGCTCCGCTTTGGGCGCTTAATGTCCTCTGAGGCCCGAGAGGGCGG 400
DB 181 GCTGTGGCACTGCTCCGCTTTGGGCGCTTAATGTCCTCTGAGGCCCGAGAGGGCGG 240
QY 401 CGGGGCTTACAGAGCGCTGCACTGCGCTTACAGAGCACTGAATATCGTCAAGGTATCCCC 460
DB 241 CGGGGCTTACAGAGCGCTGCACTGCGCTTACAGAGCACTGAATATCGTCAAGGTATCCCC 300
QY 461 GTCAGAGAGCCTTGCGCTGCTGAGCCCTTACGCGCGGCTGCCCCGCAAGCATGTG 520

```

```

DB 301 GTCCAGGAAGCCCTGCGCTGTGGAACCTTAAGCGGGCTGCCCGGACAAGCATGTG 360
QY 521 GCTGGCCCACTGAGGCTCTGCTGTGATACCAAGCTCTTAAGCCCTTTTACTGTGAGAC 580
DB 361 GCTTGGCCCACTGAGGCGCTGTGCTGTGATACCAAGCTCTTAAGCCCTTTTACTGTGAGAC 420
QY 581 CATGGGAGACATGACAGCGCTGTGCGAAGCGCCACCGTATCCCTTGAAGCTGAGGCTGCC 640
DB 421 CATGGGAGACATGACAGCGCTGTGCGAAGCGCCACCGTATCCCTTGAAGCTGAGGCTGCC 446
QY 641 GTGCTCTTCCGCAATGCGCACCGCCCTGCGCACTGTACACAGACGCTGTGCTCTG 700
DB 447 -----GACACAGCCCTGTGTGCGCACTGTGACAG----- 474
QY 701 CGTGATCTCAAGCTGTGTGCTTTGTCTTGTGCGTACCGCGTGAAGAGAGAGAGGCTGTG 760
DB 475 -----ACACGGAAGAGGCTGGGCTG 495
QY 761 GAGAACCTGAGAGACTCTGCTGTGCTGACTGTGGCCGATGATTCCTGTGGACAAGAC 820
DB 496 GAGAACCTGAGAGACTCTGCTGTGCTGACTGTGGCCGATGATTCCTGTGGACAAGAC 555
QY 821 GCGTGCCCAAGCTTACGTGGAGCTTGAATACTGAGCTCACGGGCTCATTAATCGGCAAG 880
DB 556 GCGTGCCCAAGCTTACGTGGAGCTTGAATACTGAGCTCACGGGCTCATTAATCGGCAAG 615
QY 881 GGAGCCGATGTGAGAGCTGCGGCGTGGCGCTTACCATGTGGCGGCACTACCCC 940
DB 616 GGAGCCGATGTGAGAGCTGCGGCGTGGCGCTTACCATGTGGCGGCACTACCCC 675
QY 941 TTCCAAGACTCGAGAGCTTCTGCTTTCGCGAGATCCGCCGCGGCTTACAGCCTTG 1000
DB 676 TTCCAAGACTCGAGAGCTTCTGCTTTCGCGAGATCCGCCGCGGCTTACAGCCTTG 735
QY 1001 CCTGCAAGGCTTCTGCGCCCTTGGCGCTGTGTGCTGCTGCTCTTGTGTGGAGCCA 1060
DB 736 CCTGCAAGGCTTCTGCGCCCTTGGCGCTGTGTGCTGCTGCTCTTGTGTGGAGCCA 795
QY 1061 GCTGAACGGCTCAACAGCCACAGGCACTCTGCGACCCCTGGGTGAGACAGCCGATG 1120
DB 796 GCTGAACGGCTCAACAGCCACAGGCACTCTGCGACCCCTGGGTGAGACAGCCGATG 855
QY 1121 CCTTACGCCCAACCCGATCCCATCTCTGAGAGGCTGCCAGGTGTCTCTGATGACTG 1180
DB 856 CCTTACGCCCAACCCGATCCCATCTCTGAGAGGCTGCCAGGTGTCTCTGATGACTG 915
QY 1181 GGCGTGAAGAGCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1237
DB 916 GGCGTGAAGAGCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 972
XX
RESULT 10
ABL39747
ID ABL39747 standard; cDNA, 1076 BP.
XX
XX ABL39747;
XX
DB 10-MAY-2002 (first entry)
XX
XX Human NS cDNA sequence SEQ ID NO:57.
XX
XX Human; cytostatic; osteopathic; gynaecological; neuroprotective;
XX antiinflammatory; antihistaminic; antiparasitic; ophthalmological; anti-HIV;
XX vasotropic; antiarteriosclerotic; antiinflammatory; dermatological;
XX antifibrinolytic; hypotension; antiaesthetic; immunomodulator; cardiac;
XX anticonvulsant; antidiabetic; tranquilliser; antidepressant; neuroleptic;
XX gastrointestinal; virostatic; anticancer; cerebroprotective; nootropic;
XX contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
XX endometriosis; degenerative disease; multiple sclerosis; psoriasis;
XX rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
XX inflammation; skin disorder; obesity; muscular dystrophy; AIDS;

```

KW		infertility; cardiovascular disease; coagulation disease; hypertension;
KM		ischemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
KV		diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
XX		gastric ulcer; Alzheimer's disease; gene; ss.
OS		Homo sapiens.
PN		WO200206315-A2.
XX		
PD		24-JAN-2002.
XX		
PF		17-JUL-2001; 2001WO-IL000653.
XX		
PR		18-JUL-2000; 2000IL-00137345.
XX		
FR		15-DEC-2000; 2000IL-00140354.
XX		
PA		(COMP-) COMPUGEN LTD.
PI		Mintz L, Freilich S, Bernstein J;
XX		
DR		WPI; 2002-155037/20.
XX		P-PsDB; ABB06093.
PT		One hundred and twenty eight novel nucleic acid sequences, useful for
XX		treating and diagnosing e.g. cancer, asthma and Alzheimer's.
PS		Claim 1; Page 107; 290pp; English.
XX		
CC		AB139691 to AB139618 represent novel human nucleic acid sequences
CC		encoding the proteins given in ABB06037 to ABB06164. The novel sequences
CC		(NS) can have cytoprotective, osteopathic, gynaecological, neuroprotective,
CC		antirheumatic, antiarthritic, antipsoriatic, ophthalmological, virucide,
CC		vasorelaxic, antiarteriosclerotic, antiinflammatory, dermatological,
CC		anorectic, muscular, anti-HIV, antifertility, cardiovascular,
CC		anticarcinogenic, antifibrinolytic, hypotension, antisthmatic, cardiant,
CC		immunomodulator, anticonvulsant, antidiabetic, tranquiliser, anti-ulcer,
CC		antidepressant, gastrointestinal, neurolaptic, cerebroprotective,
CC		neurotropic and contraceptive activities. The NS can be used in vaccines,
CC		gene therapy and antisense therapy. Nucleic acids, expression vectors and
CC		antibodies from the present invention can be used for treating and
CC		diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
CC		diseases, dyslexia, multiple sclerosis, rheumatoid arthritis, psoriasis,
CC		cataracts, retinitis, atherosclerosis, inflammation, skin disorders,
CC		glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
CC		disease, coagulation disease, ischaemia, hypertension, asthma, immune
CC		disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
CC		depression, schizophrenia, viral disease, gastric ulcers, stroke,
CC		Alzheimer's disease and as a contraceptive
XX		
SQ		Sequence 1076 BP; 172 A; 352 C; 346 G; 194 T; 0 U; 12 Other;
		Query Match 33.9%; Score 716.6; DB 6; Length 1076;
		Best Local Similarity 98.0%; Pred. No. 6,1e-170;
		Matches 753; Conservative 3; Mismatches 8; Indels 4; Gaps 3
QY	27	CTAGACCCCGGACAGGGCTGTGAAGCTGTGGATGCCAGTCCGACAGCGCAGGG 86
Dd	196	CTAGAACCCCGGCAAGGGCTGTGAAGCTGTGGATGCCAGTCCGACAGCGCAGG 254
QY	87	CCGGCCCACTGTCTGTGTGCTTGAGAGCTCTAAGCCCCGGCGGCGCCGACGCGG 146
Dd	255	CCGGCCCACTGTCTGTGTGCTTGAGAGCTCTAAGCCCCGGCGGCGCCGACGCGG 314
QY	147	AACGACGGGGGAGATGTGAGCACCCCCTGTGGCTGTCTTCGCGGTTCCTGTCCAGA 206
Dd	315	AACGACGGGGGAGATGTGAGCACCCCCTGTGGCTGTCTTCGCGGTTCCTGTCCAGA 374
QY	207	AGAAAGCGTTGAGATTGATGACAATTAGATACCAGAGCGTCCCGTCAGAAACGAGCTC 266
Dd	375	AGAAAGCGTTGAGATTGATGACAATTAGATACCAGAGCGTCCCGTCAGAAACGAGCTC 434
QY	267	GAAATGGGCCCCAGCCCAACTGCCTCCCTGTGCTTGGCCCTTAGACCCCACTTAATGCTC 326

Db	435	GAAAGTGGGCCCCAGCCACAGACTGACCCCCCTGCTGTGGCCCTGAGCCCACTACTGCTC	49
Qy	327	CAGATCGTGCACATGCTGTGGGCACATGCTCCCGATTGGGCGCTATGTCCTCTGGAGC	38
Db	435	CAGATCGTGCACATGCTGTGGGCACATGCTCCCGATTGGGCGCTATGTCCTCTGGAGC	554
Qy	387	CCGAGAGAGGCGGGCGGGCTTACACAGGCCCTGCACTGCTTACAGGCACTGAGTATACCT	446
Db	555	CCGAGAGAGGCGGGCGGGCTTACACAGGCCCTGCACTGCTTACAGGCACTGAGTATACCT	614
Qy	447	GCAAGCTGATCCCGTCCAGAGAGCCCTGGCCGTGTGAGCCCTTACGCGGCGTCCCC	506
Db	615	GCAAGCTGATCCCGTCCAGAGAGCCCTGGCCGTGTGAGCCCTTACGCGGCGTCCCC	674
Qy	507	CGCAACAGATGTGGCTGGGCCCATGAGAGTCTGGCTGTGATCCGAGCTTCTTAAGCCT	566
Db	675	CGCAACAGATGTGGCTGGGCCCATGAGAGTCTGGCTGTGATCCGAGCTTCTTAAGCCT	734
Qy	567	TTTTCACTCGACCCATGGGAGCATGACAGAGCTGGTGGAGACGCGCCACCGTATCCCTG	628
Db	725	TTTTCACTCGACCCATGGGAGCATGACAGAGCTGGTGGAGACGCGCCACCGTATCCCTG	794
Qy	627	AGCCTGAGGCTGCGCTGCTCTTTCGCGCAGATGAGCAACGCGCCCTGGCGCACTGTACACAGC	686
Db	795	AGCCTGAGGCTGCGCTGCTCTTTCGCGCAGATGAGCAACGCGCCCTGGCGCACTGTACACAGC	854
Qy	687	ACGGTCTGTGCTGGGTGATCTCAAGCTGTGTGCTTGTCTTTCGGTGAACCGGAGAGA	746
Db	855	ACGGTCTGTGCTGGGTGATCTCAAGCTGTGTGCTTGTCTTTCGGTGAACCGGAGAGG-	913
Qy	747	AGAACTGTGTGTGAGAACTGTGAGAGACTCCGTGGCTGTGACTGTGGGC	794
Db	914	AGAACTGTGTGTGAGAACTGTGAGAGACTCCGTGGCTGTGACTGTGGGC	959
RESULT 11			
ABL39762			
ID	ABL39762	standard; cDNA; 1076 BP.	
AC	ABL39762;		
DT	10-MAY-2002	(first entry)	
DE	Human NS cDNA sequence SEQ ID NO:12.		
XX	Human; cytosol; osteopathic; gynaecological; neuroprotective;		
KM	anti-rheumatic; anti-rheumatic; antipsoriatic; ophthalmological; anti-HIV;		
KM	vasoregic; anti-arteriosclerotic; anti-inflammatory; dermatological;		
KM	anorectic; muscular; anti-inferility; cardiovascular; anticoagulant;		
KM	anti-fibrinolytic; hypotension; antiasthmatic; immunomodulator; cardiac;		
KM	anticonvulsant; antidiabetic; tranquiliser; antidepressant; neuropoleptic;		
KM	gastrointestinal; virucide; anticancer; cerebroprotective; nocotropic;		
KM	contraepileptic; vaccine; gene therapy; cancer; osteoporosis; dystonia;		
KM	endometriosis; degenerative disease; multiple sclerosis; psoriasis;		
KM	rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;		
KM	inflammation; skin disorder; obesity; muscular dystrophy; AIDS;		
KM	infertility; cardiovascular disease; coagulation disease; hypertension;		
KM	ischemia; asthma; immune disease; epilepsy; angina; neurodegeneration;		
KM	diabetes; anxiety; depression; schizophrenia; viral disease; stroke;		
KM	gastric ulcer; Alzheimer's disease; gene; ss.		
XX	Homo sapiens.		
OS	Homo sapiens.		
PN	WO200206315-A2.		
XX	24-JAN-2002.		
PD	17-JUL-2001; 2001WO-IL000653.		
PF	18-JUL-2000; 2000IL-00137345.		
PR	15-DEC-2000; 2000IL-00140354.		
XX	(COMP-) COMPUGEN LTD.		

CC at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 2004 BP; 410 A; 563 C; 590 G; 441 T; 0 U; 0 Other;
 SQ

Query Match 29.4%; Score 621.6; DB 9; Length 2004;
 Best Local Similarity 71.9%; Pred. No. 6,4e-146;
 Matches 880; Conservativity 0; Mismatches 314; Indels 30; Gaps 4;

QY 129 GGGCCGGGGCCCAAGCGGAGACGAGCGGCGAGATGCGAGCCACCCCTCTGGCTGCTG 188
 DB 122 GGGCGAGGGCGGCTCTGAGACCTGAGACAAAGATGGAGCCACATCTCTGGCTGCTG 181
 QY 189 CGGGTTCCCTGTCTCAGAGAAAGACGGTTGAGTTGATGACACTTAAGATACGAGGTC 248
 DB 182 CGGATTTCCCTCAGAGAAAGAACCGTTGAGTTGATGACACTTAAGATGAGTGTTC 241
 QY 249 CGGTCCAGAAACGAGCTCGAAGTGGGCCCGCCAGCCAGACTGCCCCCTGCTGTTGCC 308
 DB 242 CAGTCTTAAACGAGTGAAGATGAGCTGAGCTTGAGCCACCAACCCCC----- 288
 QY 309 TTAGGCCACCTACTGCTCCAGATCTGTCACACTGCTGCTGCTGCTGCTGCTGCTG 368
 DB 289 --AGCTACCTCCCGCTCAGACTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 346
 QY 369 CCTATGCTCTCTGAGAGCCGAGAGAGGGGGGGGCTTACAGAGCCCTGCACTGGCTTA 428
 DB 347 CCTATATCTTTTGGAGACGAGAGAGCACTGCACTTATGGAGGCTGCACTGGCTTA 406
 QY 429 CAGGCACTAGATATACCTGCAAGGTGTACCCGCTCCAGAGAGCCCTGAGCTGAGAC 488
 DB 407 CAGGCACTAGATATACCTGCAAGGTGTATCCCGCCAGAGGCCAGGGGGTGTGGCAG 466
 QY 489 CCTAGCGCGCGCGCCCCCGGCAAGCACTGTGGCTGCGCCACTGAGTCTGCTGCTGTA 548
 DB 467 CATTATGCAAGACTGCTTACCAACGAGCACTGCGCCGCTCCACAGAGGTCCTGCTG 526
 QY 549 CCGAGCTCTCTACGCTTTTCACTCGAGCCATGAGGAGCAATGCAAGCTGAGTGCAG 608
 DB 527 CTCAGGTCTTATACACTTTTTCAGAGAAACCCATGAGGAGCTTTCAGACTGTGGTGC 586
 QY 609 CGCGCCACCGTATCCCTTGAAGCTGAGGCTGCTCTTCCGCGAGATGAGCCACCGCC 668
 DB 587 GCGCGCGGGTATCCCGAGCCGAGGCTGCGCGCTTTTCGAGAGATGGCTATGTGG 646
 QY 669 TGGCGCACTGTACAGAGACGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 728
 DB 647 TGGCACTGTCCACAGAGACGCGCTCATCTTTCGAGACTCAAGTTGCTGCTGCTGCT 706
 QY 729 TCGCTGACCGTGAAGAGAAAGAGTGTGTGAGAGAACTGAGAGACTGCTGCTGCTGA 788
 DB 707 TCAGCACTGTGAAGAGAGAGAGTGTGTGAGAGAACTGAGAGACTGCTGCTGCTGA 766
 QY 789 CTGGGCGAGATGATTCCTGTGTGAGCAAGACGCTGCTGCTGCTGCTGCTGCTGAG 848
 DB 767 CTGGACCAAGACGACTCTGTGTGAGCAAGACGCTGCTGCTGCTGCTGCTGCTGAG 826
 QY 849 TACTCAGCTCAGGGGCTCATCTGTGGGCAAGGCAAGCTGCTGAGGCTGCTGCTGCT 908
 DB 827 TACTCAGCTCAGGGGCTCATCTGTGGGCAAGGCAAGCTGCTGAGGCTGCTGCTGCT 886
 QY 909 CGCTTTTACAGATGTGCGCGGCTCACTACCCCTTCAGAGACTGCGAGGCTGCTGCTGCT 968
 DB 887 CGCTTTTACAGATGTGCGCGGCTCACTACCCCTTCAGAGACTGCGAGGCTGCTGCT 946
 QY 969 TCGGCAAGATCGCGCGGCTCACTACCCCTTCAGAGGCTGCTGCGGCTGCTGCTGCT 1028
 DB 947 TTGGCAAGATCGGAGAGAGAGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAT 1006
 QY 1029 GTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1088
 DB 1007 GCTTATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1066
 QY 1089 TCTGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1148

DB 1067 TGCTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1126
 QY 1149 GGGAGGCTGCCAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1208
 DB 1127 GGGAGTGTGACCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1185
 QY 1209 GAGACAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1267
 DB 1186 -----GTGGGGCTGTATGCTTACAGCCACTTAAGTGGCCCTTCACTGCAAG 1234
 QY 1268 AGTGATTTAGTTGGGGTACCTCAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1327
 DB 1235 TGTGAGTTAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1292
 QY 1328 TCAGTCTTTCAGAGAGGAGAA 1351
 DB 1293 TAAATCTCTTCTGAGAGAGAA 1316

RESULT 13
 ADB53749
 ID ADB53749 standard; DNA; 2004 BP.
 XX
 AC ADB53749;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:4291.
 XX
 KW toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
 KW toxicity marker; toxicity progression; drug screening;
 KW primary rat hepatocyte toxicity modelling; gene; ds.
 XX
 OS Rattus norvegicus.
 XX
 PN W02003065993-A2.
 PN
 PD 14-AUG-2003.
 XX
 PF 04-FEB-2003; 2003MO-US003482.
 XX
 PR 04-FEB-2002; 2002US-035171P.
 PR 13-MAR-2002; 2002US-0363534P.
 PR 08-APR-2002; 2002US-0370248P.
 PR 10-APR-2002; 2002US-0371134P.
 PR 10-APR-2002; 2002US-0371135P.
 PR 10-APR-2002; 2002US-0371150P.
 PR 11-APR-2002; 2002US-0371413P.
 PR 19-APR-2002; 2002US-0373601P.
 PR 19-APR-2002; 2002US-0373602P.
 PR 22-APR-2002; 2002US-0374139P.
 PR 08-MAY-2002; 2002US-0378370P.
 PR 09-MAY-2002; 2002US-0378652P.
 PR 09-MAY-2002; 2002US-0378653P.
 PR 09-MAY-2002; 2002US-0378655P.
 PR 09-JUL-2002; 2002US-0394233P.
 PR 09-JUL-2002; 2002US-0394250P.
 PR 04-SEP-2002; 2002US-0407688P.
 PR 28-JAN-2003; 2003US-0442900P.
 XX
 PA (GENE-) GENE LOGIC INC.
 PA
 PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
 PI Elashoff M;
 XX
 DR WPI; 2003-731472/69.
 XX
 PT Determining if a compound induces a toxic effect on a tissue or cell, for
 PT identifying hepatotoxic compounds, comprises comparing a gene expression
 PT profile of a tissue or cell sample to a database of Tox mean and non-Tox
 PT mean values.
 XX

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 29, 2004, 04:49:36 ; Search time 8342 Seconds

(without alignments)
10994.229 Million cell updates/sec

Title: US-10-070-337-16

Perfect score: 2116
Sequence: 1 ggaagcgagctccgcgcgcgt.....tatgaatacaaaaaaaaaa 2116

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: gb_ba:*
2: gb_hg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hg_hum:*
31: em_hg_inv:*
32: em_hgt_other:*
33: em_hgt_mus:*
34: em_hgt_pln:*
35: em_hgt_rod:*
36: em_hgt_mam:*
37: em_hgt_vrt:*
38: em_sy:*
39: em_hgt_hum:*
40: em_hgt_mus:*
41: em_hgt_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2116	100.0	2116	6	AX099934 Sequence
2	2100.2	99.3	2289	6	AR411699 Sequence
3	2095.4	99.0	2257	9	AK026945 Homo sapi
4	2093.8	99.0	2283	9	AK027484 Homo sapi
5	2078.2	98.2	2488	9	BC027484 Homo sapi
6	2046.2	96.7	2121	9	BC019363 Homo sapi
7	1367	64.6	153170	9	HS110367 Homo sapi
8	1072.2	50.7	1077	12	AX166518 Sequence
9	1070.6	50.6	1077	12	AX335754 Sequence
10	1069.8	50.6	1085	6	BD248318 Sequence
11	1069.2	50.5	1074	6	AR411700 Sequence
12	1053	49.8	1083	6	AX572896 Sequence
13	1053	49.8	1083	6	AF250311 Homo sapi
14	803.8	38.0	972	6	AX224734 Sequence
15	716.6	33.9	1076	6	AX364906 Sequence
16	716.6	33.9	1076	6	AX364921 Sequence
17	670.6	31.7	1280	10	MMU514260 Sequence
18	656.6	31.0	1256	10	AF358868 Homo sapi
19	621.6	29.4	2004	10	AB020967 Homo sapi
20	533.8	25.2	1969	10	BC012955 Homo sapi
21	533	25.2	541	6	AX381642 Sequence
22	505.2	23.9	2559	6	AR274998 Sequence
23	505.2	23.9	2559	9	HSU02019 Homo sapi
24	505.2	23.9	2559	6	I24432 Sequence
25	499.2	23.6	75131	2	AC023022 Homo sapi
26	354.2	16.7	396	6	AX245589 Sequence
27	318	15.0	98566	10	AI928568 Homo sapi
28	315.4	14.9	353	6	AX329832 Sequence
29	295.6	14.0	235947	2	AC133773 Homo sapi
30	295.6	14.0	282849	2	AC132555 Homo sapi
31	293.8	13.9	75131	2	AC023022 Homo sapi
32	278.2	13.1	331	11	G23310 Homo sapi
33	269.4	12.7	3319	6	AX572894 Sequence
34	269.4	12.7	3319	6	AF250310 Homo sapi
35	267.8	12.7	3317	9	AF205437 Homo sapi
36	267.8	12.7	3338	9	HSM80396 Homo sapi
37	267.8	12.7	3645	9	BC063292 Homo sapi
38	256.6	12.1	2514	10	BC006800 Homo sapi
39	253.4	12.0	1244	10	AF358866 Homo sapi
40	252.4	11.9	1682	4	AY247741 Bos tauru
41	251.4	11.9	1032	12	AY335660 Homo sapi
42	251.4	11.9	2147	9	BC002637 Homo sapi
43	251.4	11.9	2156	6	AX779968 Sequence
44	251.4	11.9	3062	9	AY245544 Homo sapi
45	251.4	11.9	4221	9	D87119 Homo sapien

ALIGNMENTS

RESULT 1	AX099934	2116 bp	DNA	linear	PAT 02-APR-2001
LOCUS	Sequence 16 from Patent WO0120004.				
DEFINITION	AX099934				
ACCESSION	AX099934.1	GI:3538944			
VERSION					
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	Yue, H., Tang, Y.T., Bandman, O., Hillman, J.L., Baughn, M.R.,				
	Azizmai, Y. and Lu, D.A.				
	Protein phosphatase and kinase proteins				

JOURNAL Patent: WO 0120004-A 16 22-MAR-2001;
Incyte Genomics, Inc. (US)
FEATURES Location/Qualifiers
source 1. 2116

```
/organism="Homo sapiens"  
/mol_type="unassigned DNA"  
/db_xref="taxon:9606"  
/note="IncYTE ID NO: 1271505CB1"
```

ORIGIN

100.0%; Score 2116; DB 6; Length 2116;

Query Match	100.0%;	Score 2116;	DB 6;	Length 2116;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2116;	Conservative 0;	Mismatches .0;	Indels 0;	Gaps 0;

QY	1	GGAGCGGCGCTCCGCGCGGTCCGCTGCTAGAACCCGCGGCAGGGCTGGAGCTGGGCTGGGA	60
Db	1	GGAGCGGCGCTCCGCGCGGTCCGCTGCTAGAGCCGGGCGAGGGCTGGAGCTGGGCTGGGA	60
QY	61	TCCGAGCTCGGACAGCAGCGCAGCGGCGGCGCCCACTGCTGGTGGCTCCGAGGCTCTGAG	120
Db	61	TCCGAGCTCGGACAGCAGCGGCGGCGGCGCCCACTGCTGGTGGCTCCGAGGCTCTGAG	120
QY	121	CCCCGGCGGCGCGCGGCGCCACGCGGAACGACGGGCGAGATGGAGGCCACCCCTCTGGC	180
Db	121	CCCCGGCGGCGCGCGGCGCCACGCGGAACGACGGGCGAGAGCGAGGCCACCCCTCTGGC	180
QY	181	TGCTCTGCGGGTTCCTGTCGAGAGAAAGGGGTGGAGTTGATGACAACTTAGATAC	240
Db	181	TGCTCTGCGGGTTCCTGTCGAGAGAAAGGGGTGGAGTTGATGACAACTTAGATAC	240
QY	241	CGAGCGTCCCGTCCAGAAACGAGCTGAGATGGGCGCCACGCCAGCTGCCCCCTGCT	300
Db	241	CGAGCGTCCCGTCCAGAAACGAGCTGAGATGGGCGCCACGCCAGCTGCCCCCTGCT	300
QY	301	GTTGGCCCTGAGGCCACCTACCTGCTCCGATGCTGTCACTGTGTGGGCCACTGGCTCCG	360
Db	301	GTTGGCCCTGAGGCCACCTACCTGCTCCGATGCTGTCACTGTGTGGGCCACTGGCTCCG	360
QY	361	TCTTGGGCGCTATGTCCTCTGAGAGCCGAGAGGGCGGCGGCGCTACAGGCGCTGCA	420
Db	361	TCTTGGGCGCTATGTCCTCTGAGAGCCGAGAGGGCGGCGGCGCTACAGGCGCTGCA	420
QY	421	CTGCCCTACAGGCACCTGAGTAACTGCAAGGTGTAACTCCCTGTCCAGAGAGCCCTGGCGT	480
Db	421	CTGCCCTACAGGCACCTGAGTAACTGCAAGGTGTAACTCCCTGTCCAGAGAGCCCTGGCGT	480
QY	481	GCTGAGCCCTACGCGCGGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT	540
Db	481	GCTGAGCCCTACGCGCGGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT	540
QY	541	GGCTGGTACCGAGCTCTCTACGCGCTTTTCACTCGAACCCATGGGGAATGACAGGCT	600
Db	541	GGCTGGTACCGAGCTCTCTCTACGCGCTTTTCACTCGAACCCATGGGGAATGACAGGCT	600
QY	601	GGTGGGAAGCGCGCACCTATCCCTAGGCTGAGGCTGCGGCTCTTCCGCGAGATGGC	660
Db	601	GGTGGGAAGCGCGCACCTATCCCTAGGCTGAGGCTGCGGCTCTTCCGCGAGATGGC	660
QY	661	CACCGCGCTGCGGCACTGTCACTCAGCAGCGGCTGTGCTTGGCTGATCTCAAGCTGTGTC	720
Db	661	CACCGCGCTGCGGCACTGTCACTCAGCAGCGGCTGTGCTTGGCTGATCTCAAGCTGTGTC	720
QY	721	CTTTGGCTTCGCTGACCGTGGAGAGAAAGCTGGTGTGGAGAACCTGAGGACTCTG	780
Db	721	CTTTGGCTTCGCTGACCGTGGAGAGAAAGCTGGTGTGGAGAACCTGAGGACTCTG	780
QY	781	CGTGTGACTGGGCGCAGATGATTCCTCTGTGGGCAAGCAGCGCGTGGCCAGCCTAGTGG	840
Db	781	CGTGTGACTGGGCGCAGATGATTCCTCTGTGGGCAAGCAGCGCGTGGCCAGCCTAGTGG	840
QY	841	ACCTGAGATCTACGCTCACGCGGCTCTCATCTGGGCAAGGCGGCGGATGTCTGAGGCT	900
Db	841	ACCTGAGATCTACGCTCACGCGGCTCTCATCTGGGCAAGGCGGCGGATGTCTGAGGCT	900

QY	901	GGGCGTGCGCTCTTCAACATGCTGGCGGCGCACTACCCCTTCCAGACCTGGAGCCTGT	960
Db	901	GGGGGTGGCGCTCTTCAACATGCTGGCGGCGCACTACCCCTTCCAGACCTGGAGCCTGT	960
QY	961	CCTGCTCTTGGGCAAGATCCGCGCGCGGAGCCTTACGACCTTGGCTCAGAGCCTCTGGAGCC	1020
Db	961	CCTGCTCTTGGGCAAGATCCGCGCGCGGAGCCTTACGACCTTGGCTCAGAGCCTCTGGAGCC	1020
QY	1021	TGCCCGCTGTCTGATTTGGCTGGCTGCTTCCTTGGTGAGAGCAGGCTTCCACACCCAC	1080
Db	1021	TGCCCGCTGTCTGATTTGGCTGGCTGCTTCCTTGGTGAGAGCAGGCTTCCACACCCAC	1080
QY	1081	AGGACATCCTCTGGACCCCTGGCGCAGAGACCCGATGGCCCTTAAAGCCCCAAC	1140
Db	1081	AGGACATCCTCTGGACCCCTGGCGCAGAGACCCGATGGCCCTTAAAGCCCCAAC	1140
QY	1141	CCATCTCTGGAGGCTGCCAGGTGCTCCATGATGACCTGGAGCTGGACGAGCCAGGA	1200
Db	1141	CCATCTCTGGAGGCTGCCAGGTGCTCCATGATGACCTGGAGCTGGACGAGCCAGGA	1200
QY	1201	AGAGGAGGGAACAGAGATGTTCGTATAGGCTAAGGACCACTTACATAGCTCAGC	1260
Db	1201	AGAGGAGGGAACAGAGATGTTCGTATAGGCTAAGGACCACTTACATAGCTCAGC	1260
QY	1261	TGCCAACAGTGAATTGATTTGGGGGTAGCTCCAGCCTTCTCCTGCTGAACTGAGC	1320
Db	1261	TGCCAACAGTGAATTGATTTGGGGGTAGCTCCAGCCTTCTCCTGCTGAACTGAGC	1320
QY	1321	CAACCTTCACTGCTCTTCCAGAGGGAAGAAAGCGAAGGCTGTGTGGAGTGTGCTGT	1380
Db	1321	CAACCTTCACTGCTCTTCCAGAGGGAAGAAAGCGAAGGCTGTGTGGAGTGTGCTGT	1380
QY	1381	ACACATCTGCTTTGTTCCACACACATGAGATCTCGCTGGGTGCTTATCAGGTGCCAAG	1440
Db	1381	ACACATCTGCTTTGTTCCACACACATGAGATCTCGCTGGGTGCTTATCAGGTGCCAAG	1440
QY	1441	CCCTGTCTCGGTGCTGGGAGTACAGCAGTGAAGCAAGAGACAAATATTCCTGCTCACA	1500
Db	1441	CCCTGTCTCGGTGCTGGGAGTACAGCAGTGAAGCAAGAGACAAATATTCCTGCTCACA	1500
QY	1501	GAGATGACAAACTGGGACCTTGTGACACTACAAACTTTTCCATGACATAGGCTCATCTGC	1560
Db	1501	GAGATGACAAACTGGGACCTTGTGACACTACAAACTTTTCCATGACATAGGCTCATCTGC	1560
QY	1561	TACACTGGGATACACTTTGTACCAAGTGTGGGCTCTCCACTGATGCTGTGCTCAGGACCTC	1620
Db	1561	TACACTGGGATACACTTTGTACCAAGTGTGGGCTCTCCACTGATGCTGTGCTCAGGACCTC	1620
QY	1621	TGTCCAAAGACATCTCCCTTTTCCAAACAAACGATGGCTTTGTATCTGTACCTTTTCA	1680
Db	1621	TGTCCAAAGACATCTCCCTTTTCCAAACAAACGATGGCTTTGTATCTGTACCTTTTCA	1680
QY	1681	GAGAAAGGAGGTATTCCTGTGSCAAAGGCTCAGGCTCTCCCTGCAACTCAGAGCCC	1740
Db	1681	GAGAAAGGAGGTATTCCTGTGSCAAAGGCTCAGGCTCTCCCTGCAACTCAGAGCCC	1740
QY	1741	AAGCCAGCTCATCTCTGGGAACTGTGTTCCAGCATCTCTGTCTCTTGAATTAAAGATTT	1800
Db	1741	AAGCCAGCTCATCTCTGGGAACTGTGTTCCAGCATCTCTGTCTCTTGAATTAAAGATTT	1800
QY	1801	CTTCCTTCCAGGCTTAAAGCTTGGGATTTGGGCGCAGAGATTAAGATCCAACTATAGGCTTA	1860
Db	1801	CTTCCTTCCAGGCTTAAAGCTTGGGATTTGGGCGCAGAGATTAAGATCCAACTATAGGCTTA	1860
QY	1861	GTTCTTGTCTAACTCAAGACTGTTCTGAATAGGCTCCAGGCTGTCAACCATGGGCT	1920
Db	1861	GTTCTTGTCTAACTCAAGACTGTTCTGAATAGGCTCCAGGCTGTCAACCATGGGCT	1920
QY	1921	TCTGACCTTACACCAAGTTTGAAGGACAGGATTTAGGACAGGGCTCTGCTGGGCCACCT	1980
Db	1921	TCTGACCTTACACCAAGTTTGAAGGACAGGATTTAGGACAGGGCTCTGCTGGGCCACCT	1980

QY 1981 GGAAGTCCAGGTGGAGCTCTTCTGGGGACAATTGGGGCCAAATCCAGGTCCATAC 2040
|
|
|
Db 1981 GGAAGTCCAGGTGGAGCTCTTCTGGGGACAATTGGGGCCAAATCCAGGTCCATAC 2040
|
|
|
QY 2041 TCTAGTTTGTGATACATGATATGTTTACCTGTGCTTAATAAGAGATTATG 2100
|
|
|
Db 2041 TCTAGTTTGTGATACATGATATGTTTACCTGTGCTTAATAAGAGATTATG 2100
|
|
|
QY 2101 AATTAATAAATAAATAA 2116
|
|
|
Db 2101 AATTAATAAATAAATAA 2116
|
|
|
RESULT 2
AR411699 2389 bp DNA linear PAT 18-DEC-2003
LOCUS AR411699
DEFINITION Sequence 7 from patent US 6638721.
ACCESSION AR411699
VERSION AR411699.1 GI:40164080
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 2389)
Meyers, R., Kapeller-Libermann, R. and Williamson, M.
AUTHORS Human protein kinases and uses therefor
TITLE Patent: US 6638721-A 7 28-OCT-2003;
JOURNAL Location/Qualifiers
FEATURES
Source 1..2389
/organism="Unknown"
/mol_type="genomic DNA"

Query Match 99.3%; Score 2100.2; DB 6; Length 2389;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2113; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 GGAAGGCGCTCCGCGCGCTCCGCTGCTAGAGCCCGGGCAGGCGTGAAGCTGGGCTGGGA 60
|
|
|
Db 222 GGAAGGCGCTCCGCGCGCTCCGCTGCTAGAGCCCGGGCAGGCGTGAAGCTGGGCTGGGA 281
|
|
|
QY 61 TCCGAGCTCGGAGCAGCAGCAGA-CGGGCGGCGCCACCTGCTGCTGAGGCTCTGA 119
|
|
|
Db 282 TCCGAGCTCGGAGCAGCAGCAGCAGGCGGCGCCACCTGCTGCTGAGGCTCTGA 341
|
|
|
QY 120 GCCCGCGGCGCGCCCGGCGCCACGCGAAGCAGCGGCGAGATGCGAGCCACCCTCTGG 179
|
|
|
Db 342 GCCCGCGGCGCGCCCGGCGCCACGCGAAGCAGCGGCGAGATGCGAGCCACCCTCTGG 401
|
|
|
QY 180 CTGCTCTCGGCGGCTTCTCTGTCAGAGAAAGCGGTTGAGTTGATGACAACTTAATA 239
|
|
|
Db 402 CTGCTCTCGGCGGCTTCTCTGTCAGAGAAAGCGGTTGAGTTGATGACAACTTAATA 461
|
|
|
QY 240 CCGAGCGTCCCGTCCAGAAAGAGCTGAGTGGGCGCCAGCCAGAGTCCCGCCCTGCC 299
|
|
|
Db 462 CCGAGCGTCCCGTCCAGAAAGAGCTGAGTGGGCGCCAGCCAGAGTCCCGCCCTGCC 521
|
|
|
QY 300 TGTGGCCCTTGAAGCCACTTACTGCTCCAGATGCTGCACTGCTGTGAGCACTGCTGCC 359
|
|
|
Db 522 TGTGGCCCTTGAAGCCACTTACTGCTCCAGATGCTGCACTGCTGTGAGCACTGCTGCC 581
|
|
|
QY 360 GTCTTGGGCGCTATGCTCTCTGAGAGCCAGAGGCGGCGGCGGCTTACAGGCGCTTGC 419
|
|
|
Db 582 GTCTTGGGCGCTATGCTCTCTGAGAGCCAGAGGCGGCGGCGGCTTACAGGCGCTTGC 641
|
|
|
QY 420 ACTGCCCTTACAGGCACTGATATATCTCAAGGTGTACCCGTCAGAGAAAGCCCTGGCG 479
|
|
|
Db 642 ACTGCCCTTACAGGCACTGATATATCTGCAAGGTGTACCCGTCAGAGAAAGCCCTGGCG 701
|
|
|
QY 480 TGTGAGAGCCCTTACAGGCGGCGGCTGCCCGCAAGAGATGTGGCTCGGCCCACTGAGGTCC 539
|
|
|
Db 702 TGTGAGAGCCCTTACAGGCGGCGGCTGCCCGCAAGAGATGTGGCTCGGCCCACTGAGGTCC 761
|
|
|

QY 540 TGGCTGTATCCAGGCTCCTCTACAGCCTTTTTCAGCTCGAGCCCATGGGAGCATGACAGCC 599
|
|
|
Db 762 TGGCTGTATCCAGGCTCCTCTACAGCCTTTTTCAGCTCGAGCCCATGGGAGCATGACAGCC 821
|
|
|
QY 600 TGGTGGAAAGCGGCCACCGTATCCTTGAAGCTTGGCGTCCGCTCTTCCGACAGATG 659
|
|
|
Db 822 TGGTGGAAAGCGGCCACCGTATCCTTGAAGCTTGGCGTCCGCTCTTCCGACAGATG 881
|
|
|
QY 660 CCAGCGCCCTGGCGGAGCTGTACAGAGCAGTGTGTCTGCTGATCTCAAGCTGTGTG 719
|
|
|
Db 882 CCAGCGCCCTGGCGGAGCTGTACAGAGCAGTGTGTCTGCTGATCTCAAGCTGTGTG 941
|
|
|
QY 720 GCTTGTCTTCTGCTGACCGTGAAGAGAAAGCTGTGTGTGAGAACCTGGAGACTCCT 779
|
|
|
Db 942 GCTTGTCTTCTGCTGACCGTGAAGAGAAAGCTGTGTGTGAGAACCTGGAGACTCCT 1001
|
|
|
QY 780 GCGTCTGACTGGGCGCAATGATTCCTGTGGAGCAAGCAGCGTCCAGCCTTACGTTG 839
|
|
|
Db 1002 GCGTCTGACTGGGCGCAATGATTCCTGTGGAGCAAGCAGCGTCCAGCCTTACGTTG 1061
|
|
|
QY 840 GACCTGAGATPACTAGCTCAGCGGCGCTGATCTCGGCGCAAGGCAAGCCGATCTGTGAGCC 899
|
|
|
Db 1062 GACCTGAGATPACTAGCTCAGCGGCGCTGATCTCGGCGCAAGGCAAGCCGATCTGTGAGCC 1121
|
|
|
QY 900 TGGGCGTGGCGCTTTCACCAATGCTGGCGGCACTACCCCTTCAGAGCTGAGAGCCTG 959
|
|
|
Db 1122 TGGGCGTGGCGCTTTCACCAATGCTGGCGGCACTACCCCTTCAGAGCTGAGAGCCTG 1181
|
|
|
QY 960 TCTGCTCTTGGGCAAGATCCGCGCGGGGCGCTTACGCTTTCAGAGCCTTTCGAGCC 1019
|
|
|
Db 1182 TCTGCTCTTGGGCAAGATCCGCGCGGGGCGCTTACGCTTTCAGAGCCTTTCGAGCC 1241
|
|
|
QY 1020 CTGCGCGCTGCTGGGTTGCTGCGCTCTTCTCGCGGAGCCAGCTGAAGGCTCACAGCA 1079
|
|
|
Db 1242 CTGCGCGCTGCTGGGTTGCTGCGCTCTTCTCGCGGAGCCAGCTGAAGGCTCACAGCA 1301
|
|
|
QY 1080 CAGGCACTCTCTGACCCCTGCTGGCGAGCAGCCAGATCCCTTACCCCAACCCGAT 1139
|
|
|
Db 1302 CAGGCACTCTCTGACCCCTGCTGGCGAGCAGCCAGATCCCTTACCCCAACCCGAT 1361
|
|
|
QY 1140 CCGATCTCTGGAGAGCTGCGCCAGGTGTCCTGATGGACTGGGCGTGGAGCAAGGCAAGG 1199
|
|
|
Db 1362 CCGATCTCTGGAGAGCTGCGCCAGGTGTCCTGATGGACTGGGCGTGGAGCAAGGCAAGG 1421
|
|
|
QY 1200 AAGAGGAGGAGACAGAGAGTGTCTGATATGCTTGAAGCAACCTTAAGCTCTAG 1259
|
|
|
Db 1422 AAGAGGAGGAGACAGAGAGTGTCTGATATGCTTGAAGCAACCTTAAGCTCTAG 1481
|
|
|
QY 1260 CTGCCAAGTGAATGATTTGGGGGTAGCTCCAGAGCTTCTCTGCGCTCTGAAGCTGAG 1319
|
|
|
Db 1482 CTGCCAAGTGAATGATTTGGGGGTAGCTCCAGAGCTTCTCTGCGCTCTGAAGCTGAG 1541
|
|
|
QY 1320 CCAAACTTCAATGCTCTTCCAGAAAGGAGAAAGCAGAGCTGTGAGATGCTGTG 1379
|
|
|
Db 1542 CCAAACTTCAATGCTCTTCCAGAAAGGAGAAAGCAGAGCTGTGAGATGCTGTG 1601
|
|
|
QY 1380 TACACATCTGCTTTGTTCCACACACAGAGTCTCTGCTGGGCTTATATAGTGGCA 1439
|
|
|
Db 1602 TACACATCTGCTTTGTTCCACACACAGAGTCTCTGCTGGGCTTATATAGTGGCA 1661
|
|
|
QY 1440 GCCCTGTTCTCGGTGCTGGAGTACAGCAGTGAAGCAAGAGCAATATTCCTGCTCAC 1499
|
|
|
Db 1662 GCCCTGTTCTCGGTGCTGGAGTACAGCAGTGAAGCAAGAGCAATATTCCTGCTCAC 1721
|
|
|
QY 1500 AGAGATACAAACTGGGATCTTGAAGTGAACAACCTTTTCATGACCATAGTACTGT 1559
|
|
|
Db 1722 AGAGATACAAACTGGGATCTTGAAGTGAACAACCTTTTCATGACCATAGTACTGT 1781
|
|
|
QY 1560 CTACACTGGGATCACTTTGTAACAGTGTGGGCTCCACATGATGCTGTGCTCAGGAGCCT 1619
|
|
|
Db 1782 CTACACTGGGATCACTTTGTAACAGTGTGGGCTCCACATGATGCTGTGCTCAGGAGCCT 1841
|
|
|
QY 1620 CTGTCCAAGGACATCCCTTTTCAAAACAAACAGCTGCTTGTATCTTGTACTCTTTC 1679
|
|
|

Db 1842 CTGTCCAGGCAATCCCTTTCACAAACACAGCTCCCTTTGATCTTGTACTTTTC 1901
 QY 1880 AGAGAAAGGAGATATCCCTGTGCAAGAGGCTCCAGGCTCTCCCTGCAACTCAGAGC 1739
 Db 1902 AGAGAAAGGAGATATCCCTGTGCAAGAGGCTCCAGGCTCTCCCTGCAACTCAGAGC 1961
 QY 1740 CAAGCCGAGCTCATCTGTGGAACTGTGTCCAGATCTCTGTCTCTCTTTGAATAGAT 1799
 Db 1962 CAAGCCGAGCTCATCTGTGGAACTGTGTCCAGATCTCTGTCTCTTTGAATAGAT 2021
 QY 1800 TCTGCTTCAGGCTCAAGCTGTGGATTTGGGCGAGATTAAGAACTCAAACTAGAGCT 1859
 Db 2022 TCTGCTTCAGGCTCAAGCTGTGGATTTGGGCGAGATTAAGAACTCAAACTAGAGCT 2081
 QY 1860 AGTCTTGTCTAATCAAGACTGTTCTGGAATGAGGGTCCAGGCTGTCAACCATGAGGC 1919
 Db 2082 AGTCTTGTCTAATCAAGACTGTTCTGGAATGAGGGTCCAGGCTGTCAACCATGAGGC 2141
 QY 1920 TTCTGACCTGAGCACCAGGTTGAGAGGACAGATTAAGCAGGCTGTCTGTGCGCACC 1979
 Db 2142 TTCTGACCTGAGCACCAGGTTGAGAGGACAGATTAAGCAGGCTGTCTGTGCGCACC 2201
 QY 1980 TGGAAAGTCCGAGGAGCTCTTCTGAGGACACTTGGGGTCCACAACTCCAGTCCATA 2039
 Db 2202 TGGAAAGTCCGAGGAGCTCTTCTGAGGACACTTGGGGTCCACAACTCCAGTCCATA 2261
 QY 2040 CTCTAGTCTTGGATACATGATGATGTTTACTCTGTGCTTAATTAAGAGATTAAT 2099
 Db 2262 CTCTAGTCTTGGATACATGATGATGTTTACTCTGTGCTTAATTAAGAGATTAAT 2321
 QY 2100 GAAATTAATAAAAAA 2116
 Db 2322 GAAATTAATAAAAAA 2338

RESULT 3
 AK026945
 LOCUS Homo sapiens cDNA: FLJ23292 fls, clone HEP10334.
 DEFINITION AK026945
 ACCESSION AK026945.1 GI:10439923
 VERSION oligo capping: fls (full insert sequence).
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCES
 AUTHORS Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y.,
 Okamoto, S., Okitani, R., Ota, T., Suzuki, Y., Obayashi, M., Nishi, T.,
 Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.
 TITLE NEDO human cDNA sequencing project
 JOURNAL Unpublished
 AUTHORS 2 (bases 1 to 2257)
 TITLE Direct Submission
 JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
 University of Tokyo, Laboratory of Genome Structure Analysis, Human
 Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
 Japan (E-mail: f1cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
 Fax:81-3-5449-5416)
 COMMENT NEDO human cDNA sequencing project supported by Ministry of
 International Trade and Industry of Japan. cDNA full insert
 sequencing: Research Association for Biotechnology; cDNA library
 construction: 5'- & 3'-end one pass sequencing: Department of
 Virology and Human Genome Center, Institute of Medical Science,
 University of Tokyo (partly supported by Science and Technology
 Agency).
 FEATURES
 source Location/Qualifiers
 1..2257
 /organism="Homo sapiens"
 /mol_type="mRNA"

/db xref="taxon:9606"
 /clone="HEP10334"
 /cell_line="HEP2"
 /cell_type="hepatoma"
 /clone_id="HEP"
 /note="Cloning vector pHE18FLJ3"
 295..1371
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAB15597.1"
 /db_xref="GI:10439924"

ORIGIN

Query Match 99.0%; Score 2095.4; DB 9; Length 2257;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 2110; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

1 GAGAGCGGCTCCGCGCGCTCGCTGTAGAGACCCGGGAGGCTGAGCTGGGCTGGGA 60
 134 GAGAGCGGCTCCGCGCGCTCGCTGTAGAGACCCGGGAGGCTGAGCTGGGCTGGGA 193
 61 TCCGAGCTCGGACAGCAGCA -GGGCGCGGCGCACTGTGTGCTGCTGAGCTGGA 119
 194 TCCGAGCTCGGACAGCAGCA -GGGCGCGGCGCACTGTGTGCTGCTGAGCTGGA 253
 120 GCCCGCGGCGCGCGGCGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 179
 254 GCCCGCGGCGCGCGGCGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 313
 180 CTGCTCTGCGGCTTCCCTGTCTGAGAGAGACCGGTTGAGTTGATGACAACTTAGATA 239
 314 CTGCTCTGCGGCTTCCCTGTCTGAGAGAGACCGGTTGAGTTGATGACAACTTAGATA 373
 240 CCGAGCGTCCGCTCCAGAAAGAGGCTCGAAGTGGGCGGCGGCGGCGGCGGCGGCGGCGG 299
 374 CCGAGCGTCCGCTCCAGAAAGAGGCTCGAAGTGGGCGGCGGCGGCGGCGGCGGCGGCGG 433
 300 TGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 359
 434 TGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 493
 360 GTCTTGGGCGGCTATGCTCTCTGAGAGCCCGGAGAGGCGGCGGCGGCGGCGGCGGCGG 419
 494 GTCTTGGGCGGCTATGCTCTCTGAGAGCCCGGAGAGGCGGCGGCGGCGGCGGCGGCGG 553
 420 ACTGCGCTTACAGGCACTGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 479
 554 ACTGCGCTTACAGGCACTGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 613
 480 TGTGAGGCGCTTACAGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 539
 614 TGTGAGGCGCTTACAGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 673
 540 TGTGCTGTACCAAGCTCTCTTACGCTTTTTCACCTGAGAGCCCATGAGGACAGAGCC 599
 674 TGTGCTGTACCAAGCTCTCTTACGCTTTTTCACCTGAGAGCCCATGAGGACAGAGCC 733
 600 TGTGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 659
 734 TGTGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 793
 660 CCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 719
 794 CCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 853
 720 GCTTTGTCTTGTGCTACCGTGAAGAGAAAGCTGCTGCTGAGAAACCTGGAAGACTCTCT 779

```

Db      854 GCTTTGCTCTTCGCTGACCGCTGAGAGAGAAAGAGTGTGCTGAGAACTTGAGAGACTCTCT 913
Qy      780 GGGTGTGACTGAGGCGCAGATGATTCCTGTGGACAAAGCAGCGGTGCCAGGCTTAAGTGG 839
Db      914 GGGTGTGACTGAGGCGCAGATGATTCCTGTGGACAAAGCAGCGGTGCCAGGCTTAAGTGG 973
Qy      840 GACCTGAGATACCTGAGCTCAGCGGCTCTACTCTGGGAGCAAGGACGCTGCTGTGGAGCC 899
Db      974 GACCTGAGATACCTGAGCTCAGCGGCTCTACTCTGGGAGCAAGGACGCTGCTGTGGAGCC 1033
Qy      900 TGGGGGTGGCGCTCTTCAACCATGCTGGCGGCGCACTACCCCTTCCAGAGACTGGAGGCTGG 959
Db      1034 TGGGGGTGGCGCTCTTCAACCATGCTGGCGGCGCACTACCCCTTCCAGAGACTGGAGGCTGG 1093
Qy      960 TCTGTCTCTTGGGCAAGATCCGCGGCGGCGCTAGCGCTTGGCTGAGCGCTCTGCGGCC 1019
Db      1094 TCTGTCTCTTGGGCAAGATCCGCGGCGGCGCTAGCGCTTGGCTGAGCGCTCTGCGGCC 1153
Qy      1020 CTGCGCGGCTGTGGTGTGGCTGCTCTCTTCTGTGGGAGCGAGCTGAAGGCTTCAAGCCA 1079
Db      1154 CTGCGCGGCTGTGGTGTGGCTGCTCTCTTCTGTGGGAGCGAGCTGAAGGCTTCAAGCCA 1213
Qy      1080 CAGGATCTCTCTGACACCCCTGCTGCGGACAGAGACCCGATGCTTGAAGCCCAAGCCGAT 1139
Db      1214 CAGGATCTCTCTGACACCCCTGCTGCGGACAGAGACCCGATGCTTGAAGCCCAAGCCGAT 1273
Qy      1140 CCGATCTCTGAGGAGCTGCGGAGGTGTCTCTGATGACCTGAGGCTGAGCGAAGCCAGG 1199
Db      1274 CCGATCTCTGAGGAGCTGCGGAGGTGTCTCTGATGATGCTGAGGCTGAGCGAAGCCAGG 1333
Qy      1200 AAGAGAGAGGAGAGACAGAGAAAGTGTCTGTATGCTAGAGCAACCTTACTACAGCTCAG 1259
Db      1334 AAGAGAGAGGAGAGACAGAGAAAGTGTCTGTATGCTAGAGCAACCTTACTACAGCTCAG 1393
Qy      1260 CTGCAACAGTGGATTGAGTTGGGGGAGTAGCTCCAAAGCCTTCTCGGCTCTGAACTAG 1319
Db      1394 CTGCAACAGTGGATTGAGTTGGGGGAGTAGCTCCAAAGCCTTCTCGGCTCTGAACTAG 1453
Qy      1330 CCAAACTTTCAGTGTCTTCCAGAGAGGAGAAAGGAGAAAGCCTGTGTGAGTGTGCTGTG 1379
Db      1454 CCAAACTTTCAGTGTCTTCCAGAGAGGAGAAAGGAGAAAGCCTGTGTGAGTGTGCTGTG 1513
Qy      1380 TACAATCTGCTTGTGTCACACACATGCAAGTCTCTGCTGGGTCTTATCAGGTGCCAA 1439
Db      1514 TACAATCTGCTTGTGTCACACACATGCAAGTCTCTGCTGGGTCTTATCAGGTGCCAA 1573
Qy      1440 GCCCTGTTCTCGGTGCTGGAGTACAGACAGTGAAGAGAGCAATATTCCTGCTCAG 1499
Db      1574 GCCCTGTTCTCGGTGCTGGAGTACAGACAGTGAAGAGAGCAATATTCCTGCTCAG 1633
Qy      1500 AAGATGACAAACCTGGCATCTTGAAGCTGACAAACATTTTCCATGACATAGTCACTGT 1559
Db      1634 AAGATGACAAACCTGGCATCTTGAAGCTGACAAACATTTTCCATGACATAGTCACTGT 1693
Qy      1560 CTACACTGGGTGACACTTGTGACAGATGTGCGCTTCACTGATGCTGTGCTCAGGCACT 1619
Db      1694 CTACACTGGGTGACACTTGTGACAGATGTGCGCTTCACTGATGCTGTGCTCAGGCACT 1753
Qy      1620 CTGTCTCAAGAGCAATCCCTTTCACAAACAAACAGCTGCTTGTGATCTTGACCTTTC 1679
Db      1754 CTGTCTCAAGAGCAATCCCTTTCACAAACAAACAGCTGCTTGTGATCTTGACCTTTC 1813
Qy      1680 AAGAGAAAGGAGGATCCCTGTGCGCAAGGCTCAGAGGCTCTCCCTGCAACTCAGAGACC 1739
Db      1814 AAGAGAAAGGAGGATCCCTGTGCGCAAGGCTCAGAGGCTCTCCCTGCAACTCAGAGACC 1873
Qy      1740 CAAGCCAGCTCACTCTGGGAAAGTGTGCTCCAGAGATCTGTGCTCTTGAATTAAGAGAT 1799
Db      1874 CAAGCCAGCTCACTCTGGGAAAGTGTGCTCCAGAGATCTGTGCTCTTGAATTAAGAGAT 1933
Qy      1800 TCTCTCTTCAAGCCCTTAAGCCTTGAGATTTGGGCAAGAGATTAAGAACTAATAGAGCT 1859

```

```

Db      1934 TCTCTTCCAGGCTTAAGCCTTGAGATTGGGCGCAGAGATTAAGATCAAACTATAGAGCT 1993
Qy      1860 AGTTCTTGTCTAATCAAGACTGTCTGGAATGAGGAGTCCAGGCTGTCAACCATAGGAGC 1919
Db      1994 AGTTCTTGTCTAATCAAGACTGTCTGGAATGAGGAGTCCAGGCTGTCAACCATAGGAGC 2053
Qy      1920 TTCTGACCTGAGCACCAGAGTTGAGGAGCAGAGATTAGGAGGAGTCTGTCTGTGGCACC 1979
Db      2054 TTCTGACCTGAGCACCAGAGTTGAGGAGCAGAGATTAGGAGGAGTCTGTCTGTGGCACC 2113
Qy      1980 TGGAAAGTCCAGAGTGGAGACTTTCTGGGAGACACTTGGGAGTCAACAATCCAGGTCCATA 2039
Db      2114 TGGAAAGTCCAGAGTGGAGACTTTCTGGGAGACACTTGGGAGTCAACAATCCAGGTCCATA 2173
Qy      2040 CTCTAGTGTGATGATACCATGAGTATGATGTTTACCTGTGCTGCTTAAGAGAGATTAT 2099
Db      2174 CTCTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2233
Qy      2100 GAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2116
Db      2234 GAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2250

RESULT 4
BC027484
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2283)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heist,F.,
Diatchenko,L., Marsina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schneitz,T.E., Brownstein,M.J., Udell,T.B., Toshilyuk,I.S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulik,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Faney,J., Helton,E., Kelleman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,X.S., Krzywicki,M.I., Skalska,U., Smalins,D.E.,
Schererch,A., Schein,J.B., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL
MEDLINE
PUBMED
1247932
2 (bases 1 to 2283)
Strausberg,R.
Direct Submission
Submitted (04-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
COMMENT
NIH-MGC Project URL: http://mgc.ncl.nih.gov
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILMIL)
DNA Sequencing by: Baylor College of Medicine Human Genome

```

Sequencing Center
Center code: BCM-HSNC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunnarone, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louised, H.,
Kovits, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
Series: IRAX Plates: 52 Row: f Column: 11
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.

FEATURES

source
1..2283
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:34909 IMAGE:5104452"
/issue_type="Cervix, carcinoma"
/clone_1ib="NH MGC_12"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
1..2283
/gene="C20orf97"
/note="Synonyms: d11103g7.3, SKIP3, TRB3, SINK"
/db_xref="LocustID:57761"
295..1371
/codon_start=1
/product="chromosome 20 open reading frame 97"
/protein_id="AAH27484.1"
/db_xref="GI:20071611"
/translation="MKATPLAAPAGSLSRKKRLIEDNDLTERPVQKARASGPORLP
PCLLPLSEPTADRAATAVATASRLGPYVLEPEEGRAARALHCPTEYTKVAVQ
EALVALEPARLPKPHVAPTEVLAVGLLAFPTRTGDMHSVSRHLPPEEAA
VIFROMATLAHCHOHGIVIRDLKLCRFADRRKKVLEMLDSCVLTGPDLSMD
KACRAVYPELISLSASYSGRADVWSLGVALLFTMLAGHYFOSEPTVLTKGKRRG
AALPAGLSAPARCLRLRREPARPRLTATLILHPLRQDPMPLPRLSHLEAAQ
VVPDGLDEABEEBGRVLYG"
532..1239
/note="S.TK; Region: Serine/Threonine protein kinases,
catalytic domain. Phosphotransferases of the serine or
threonine-specific kinase subfamily. The enzymatic
activity of these protein kinases is controlled by
phosphorylation of specific residues in the activation
segment of the catalytic domain, sometimes combined with
reversible conformational changes in the C-terminal
autoregulatory tail"
/db_xref="CDD:c00180"

ORIGIN

Query Match 99.0%; Score 2093.8; DB 9; Length 2283;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2109; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
1 GGAAGCGGCTCCGCGGCGTCCGCTGCTAGAGACCCGGGCGAGGCTGAGCTGGGA 60
134 GGAAGCGGCTCCGCGGCGTCCGCTGCTAGAGACCCGGGCGAGGCTGAGCTGGGA 193
61 TCCCGAGCTCCGAGAGAGGCA-CGGGCGGGCCCACTGCTGCTGAGAGGCTCTGA 119
194 TCCCGAGCTCCGAGAGAGGCA-CGGGCGGGCCCACTGCTGCTGAGAGGCTCTGA 253
120 GCCCGCGGCGGCGGCGGCGGCA-CGAGGCGGCGAGATGCGAGCCACTCTGG 179
254 GCCCGCGGCGGCGGCGGCGGCA-CGAGGCGGCGAGATGCGAGCCACTCTGG 313
180 CTGCTCTCTGGGTTCCCTCTCCAGAGAAAGGCTTGGAGTTGGATGACAACTTAGTA 239
314 CTGCTCTCTGGGTTCCCTCTCCAGAGAAAGGCTTGGAGTTGGATGACAACTTAGTA 373

240 CCGAGCGTCCCGTCCAGAAAGAGCTCGAAGTGGGCCCCAGCCAGACTGCCCCCTGCC 299
374 CCGAGCGTCCCGTCCAGAAAGAGCTCGAAGTGGGCCCCAGCCAGACTGCCCCCTGCC 433
300 TGTGGCCCTGAGGCCACCTACTGCTCGAGATGCTGCAATGCTGTGACACTGCTGCC 359
434 TGTGGCCCTGAGGCCACCTACTGCTCGAGATGCTGCAATGCTGTGACACTGCTGCC 493
360 GTCTTGGGCGCTTATGCTCTCTGAGAGCCGAGAGAGGGGGGGGCTTACAGAGGCTTGC 419
494 GTCTTGGGCGCTTATGCTCTCTGAGAGCCGAGAGAGGGGGGGGCTTACAGAGGCTTGC 553
420 ACTGCCCTACAGGCACTAGATATACCTGCAAGTGTATCCCGCTCCAGAGAGGCTTGC 479
554 ACTGCCCTACAGGCACTAGATATACCTGCAAGTGTATCCCGCTCCAGAGAGGCTTGC 613
480 TGTGAGAGCCCTTACGCGGCGGCTGCCCCGACAGCATGTGCTTGGCCCACTGAGGTTC 539
614 TGTGAGAGCCCTTACGCGGCGGCTGCCCCGACAGCATGTGCTTGGCCCACTGAGGTTC 673
540 TGGCTGTATCCAGAGCTCTCTTACGCTTTTTCACCTGAGACCATGAGGACATGACAGCC 599
674 TGGCTGTATCCAGAGCTCTCTTACGCTTTTTCACCTGAGACCATGAGGACATGACAGCC 733
600 TGTGCGAAAGCGCCACCGATATCCCTGAGCCTGAGGCTGCGTCTCTTCCGCAAGTGG 659
734 TGTGCGAAAGCGCCACCGATATCCCTGAGCCTGAGGCTGCGTCTCTTCCGCAAGTGG 793
660 CCAAGCGGCTGGCGGCACTGTACCGACAGCGTGTGCTGTGATCTCAAGCTGTGTC 719
794 CCAAGCGGCTGGCGGCACTGTACCGACAGCGTGTGCTGTGATCTCAAGCTGTGTC 853
720 GCTTGTGCTGCTGACCGGTGAGAGAGAGAGCTGCTGAGAGAACTGAGAACTGCTCT 779
854 GCTTGTGCTGCTGACCGGTGAGAGAGAGAGCTGCTGAGAGAACTGAGAACTGCTCT 913
780 GCGTGTGCTGCTGAGGCGAGATGATTCCTGTGGAGAACGACGCGTGCCTTACGCTG 839
914 GCGTGTGCTGCTGAGGCGAGATGATTCCTGTGGAGAACGACGCGTGCCTTACGCTG 973
840 GACCTGAGATATCTAGCTTACGAGGCTCTTATCTCGGGCAAGCAAGCCGATGTCTGAGGC 899
974 GACCTGAGATATCTAGCTTACGAGGCTCTTATCTCGGGCAAGCAAGCCGATGTCTGAGGC 1033
900 TGGGGGTGGCGCTCTTACCATGCTGAGCGGCGCACTACCTTCCAGAGACTCGAGCTG 959
1034 TGGGGGTGGCGCTCTTACCATGCTGAGCGGCGCACTACCTTCCAGAGACTCGAGCTG 1093
960 TCTGTCTTTCGCAAGATCCGCGGCGGCTTACGCTTTCGCTGAGGCTCTTTCGCGCC 1019
1094 TCTGTCTTTCGCAAGATCCGCGGCGGCTTACGCTTTCGCTGAGGCTCTTTCGCGCC 1153
1020 CTGCGCGCTGTCTGCTGCTGCTGCTCTCTCTGAGAGCACTGAGAGGCTTACAGGCCA 1079
1154 CTGCGCGCTGTCTGCTGCTGCTGCTCTCTCTGAGAGCACTGAGAGGCTTACAGGCCA 1213
1080 CAGGCACTCTCTGCAACCCCTGCTGCAAGAGCCGATGCTTACGCTTACGCTTACGCTT 1139
1214 CAGGCACTCTCTGCAACCCCTGCTGCAAGAGCCGATGCTTACGCTTACGCTTACGCTT 1273
1140 CCGATCTTTCGAGAGCTGCGCAAGTGTCTCTGATGACTGAGGCTGAGAGGCAAGG 1199
1274 CCGATCTTTCGAGAGCTGCGCAAGTGTCTCTGATGACTGAGGCTGAGAGGCAAGG 1333
1200 AAGAGAGAGAGAGAGAGAGAGTGTCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1259
1334 AAGAGAGAGAGAGAGAGAGAGTGTCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1393
1260 CTGCGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1319
1394 CTGCGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1453
1320 CCAAGCTTTCAGTCTTTCAG 1379

Db 1100 GCTTGTCTTCGCTGACCGTGAAGAGAAAGCTGGTGTGTAAGAACTGTGAAGACTCTGAGGACTCCT 1159
 QY 780 GCGTGTGATTTGGGCGGAGATGATTTCCCTGTGGGACAAAGCAACGCTGCTCCAGCTTACGTGG 839
 Db 1160 GCGTGTGATTTGGGCGGAGATGATTTCCCTGTGGGACAAAGCAACGCTGCTCCAGCTTACGTGG 1219
 QY 840 GACCTGAGATACAGCTCAAGGAGCTCATATCTGGGCAAGGACCGATGTCTGAGCC 899
 Db 1220 GACCTGAGATACAGCTCAAGGAGCTCATATCTGGGCAAGGACCGATGTCTGAGCC 1279
 QY 900 TGGGCGTGGGCTCTTACCATGTGGCGGACCACTACCCCTTTCAGAGACTCGAGCTG 959
 Db 1280 TGGGCGTGGGCTCTTACCATGTGGCGGACCACTACCCCTTTCAGAGACTCGAGCTG 1339
 QY 960 TCCGTCTTGGGCAAGATCCGCGGGGCTACCGCTTGTCTGAGGCTCTCGGCC 1019
 Db 1340 TCTGTCTTGGGCAAGATCCGCGGGGCTACCGCTTGTCTGAGGCTCTCGGCC 1399
 QY 1020 CTGCGCGCTGTGTGTGCTGCTCTCTGTGTGGGAGCCAGCTGAACGCTCAAGCCA 1079
 Db 1400 CTGCGCGCTGTGTGTGCTGCTCTCTGTGTGGGAGCCAGCTGAACGCTCAAGCCA 1459
 QY 1080 CAGGCACTCTTCTGCAACCCCTGCTGCAAGAACCCGATGCCCTTAAGCCCAACCCGAT 1139
 Db 1460 CAGGCACTCTTCTGCAACCCCTGCTGCAAGAACCCGATGCCCTTAAGCCCAACCCGAT 1519
 QY 1140 CCATCTCTGGGAGGCTGCCAGTGGTCCCTGATGGAAGGAGGCTGGAAGAACGAGG 1199
 Db 1520 CCATCTCTGGGAGGCTGCCAGTGGTCCCTGATGGAAGGAGGCTGGAAGAACGAGG 1579
 QY 1200 AAGAGAGGAGAGACAGAGAGTGTCTGTATGCTAGAGCAACCCCTACTACAGCTCAG 1259
 Db 1580 AAGAGAGGAGAGACAGAGAGTGTCTGTATGCTAGAGCAACCCCTACTACAGCTCAG 1639
 QY 1260 CTGCGCAACAGTGAATTTGGGGGTAGCTCCAAAGCTTCTCTGCTCTGAACTGAG 1319
 Db 1640 CTGCGCAACAGTGAATTTGGGGGTAGCTCCAAAGCTTCTCTGCTCTGAACTGAG 1699
 QY 1320 CCAAACCTTCAGTGCCTTCAGAGAGGAGAGAGCAAGCCGTGTGAGAGTGTGCTG 1379
 Db 1700 CCAAACCTTCAGTGCCTTCAGAGAGGAGAGAGCAAGCCGTGTGAGAGTGTGCTG 1759
 QY 1380 TACACATCTGCTTGTTCACACACATGACATGCTGCTGGTGTGCTTACAGGTGCCAA 1439
 Db 1760 TACACATCTGCTTGTTCACACACATGACATGCTGCTGGTGTGCTTACAGGTGCCAA 1819
 QY 1440 GCGCTGTCTCGGTGCTGGAGTACAGCACTGAGCAAGAGCAATATTCCTGCTCAC 1499
 Db 1820 GCGCTGTCTCGGTGCTGGAGTACAGCACTGAGCAAGAGCAATATTCCTGCTCAC 1879
 QY 1500 AGAGATGACAAATCTGGCATCTTGAAGTGAACAACCTTTCATGACCAATAGGTCACTGT 1559
 Db 1880 AGAGATGACAAATCTGGCATCTTGAAGTGAACAACCTTTCATGACCAATAGGTCACTGT 1939
 QY 1560 CTACACTGGGTACACTTGTACAGAGTGTGCGCTTCCACTGATGCTGGTGTCAAGGACCT 1619
 Db 1940 CTACACTGGGTACACTTGTGTACAGAGTGTGCGCTTCCACTGATGCTGGTGTCAAGGACCT 1999
 QY 1620 CTGTCCAAGGACATCCCTTTTCAAAACCAACAGCTGCTTGTATCTTGTACCTTTTC 1679
 Db 2000 CTGTCCAAGGACATCCCTTTTCAAAACCAACAGCTGCTTGTATCTTGTACCTTTTC 2059
 QY 1680 AGAGAAAGGAGGTATCCCTGTGCAAAAGGCTCCAGGCTCTCCCTGCAACTCAGGACC 1739
 Db 2060 AGAGAAAGGAGGTATCCCTGTGCAAAAGGCTCCAGGCTCTCCCTGCAACTCAGGACC 2119
 QY 1740 CAAGCCAGCTCACTCTGGAGACTGTTCACAGCACTTCTGTCTTGTATTAAGAT 1799
 Db 2120 CAAGCCAGCTCACTCTGGAGACTGTTCACAGCACTTCTGTCTTGTATTAAGAT 2179
 QY 1800 TCTCTTTCAGGCTTAAGCTTGGAGATTTGGGCTAGAGATTAAGATCAAACTATGAGGCT 1859
 Db 2180 TCTCTTTCAGGCTTAAGCTTGGAGATTTGGGCTAGAGATTAAGATCAAACTATGAGGCT 2239

QY 1860 AGTCTTGTCTAATCTCAAGACTGTCTCTGAAATGAGGGTCCAGGCTGTCAACCATGGGGC 1919
 Db 2240 AGTCTTGTCTAATCTCAAGACTGTCTCTGAAATGAGGGTCCAGGCTGTCAACCATGGGGC 2299
 QY 1920 TTCTGACCTTGAAGACCAAGGTTGAGGAGCAGAGATTAGGACAGGGTCTGTCTGTGGCACC 1979
 Db 2300 TTCTGACCTTGAAGACCAAGGTTGAGGAGCAGAGATTAGGACAGGGTCTGTCTGTGGCACC 2359
 QY 1980 TGGAAAGTCCAGGTGGAGCTTCTTGGGAGCACTTGGGGTCCAAATCCAGGTCCATA 2039
 Db 2360 TGGAAAGTCCAGGTGGAGCTTCTTGGGAGCACTTGGGGTCCAAATCCAGGTCCATA 2419
 QY 2040 CTCTAGTTTGGATACCATGATATGATATGTTTACTGTGCTTAAAGAGAAATTAT 2099
 Db 2420 CTCTAGTTTGGATACCATGATATGATATGTTTACTGTGCTTAAAGAGAAATTAT 2479
 QY 2100 GAAATTA 2106
 Db 2480 GAAATTA 2486

RESULT 6
 BC019363
 LOCUS
 DEFINITION Homo sapiens chromosome 20 open reading frame 97, mRNA (cDNA clone
 ACCESION BC019363
 VERSION BC019363.1 GI:17939414
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS
 1 (bases 1 to 2121)
 Klausner, R.D., Collins, F.S., Wagner, L.H., Grouse, L.H., Derge, J.G.,
 Altschul, S.F., Zeeberg, B., Burow, K.H., Scheffer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullah, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hu, X., Gibbs, R.A.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallus, D.E.,
 Schermer, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 2 (bases 1 to 2121)
 REFERENCE
 AUTHORS
 DIRECT SUBMISSION
 JOURNAL

REMARK
 COMMENT
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E.B. Consortium (LNLML)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada

info@cgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prahbu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smalnu, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNl at: <http://image.llnl.gov>
Series: IRAL Plate: 8 Row: 5 Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 11056039.

FEATURES

source

1..2121

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:854 IMAGE:3528490"

/issue_type="Muscle, rhabdomyosarcoma"

/clone_1b="NIH MGC 17"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

1..2121

/gene="C20orf97"

/note="synonyms: dJ1103g7.3, SKIP3, TRB3, SINK"

/db_xref="LocustID:57761"

108..1184

/codon_start=1

/product="chromosome 20 open reading frame 97"

/note="protein_id="AAH19363.1"

/db_xref="GI:17939415"

/db_xref="LocustID:57761"

/translation="MRAPPLAPAGSLSKRRRLDLNDLTERPVOKRARGPRLP
PCLPLSPPTADRAATAVAATSRIGPVYLLRPEBGRALQALHCTGTYTKYVQ
EALAVIEPARLPFHNVARPTFVLAIGTOLVAFTRHGDHSLVRSNRLPEBPA
VLFRQWATLALAHQHOVLRLKLCRFVADREKRLVLEMLDSCVLTGDDSLMD
KACPAVYGEIILSSRAYSGRADWVSLGVALFTMLAGHYFPDSEPVILFGKLRG
AALPAGLSAPARCIVRCILREPARRLTATGILLHPMLRQDPMPLAPRSLWEEAQ
VYVDGLDEADEEBEBSDDREVLYYG"

345..1052

/note="S_TKc; Region: Serine/Threonine protein kinases,
catalytic domain"

/db_xref="CD:smart00220"

misc_feature

ORIGIN

Query Match

Best Local Similarity 96.7%; Score 2046.2; DB 9; Length 2121;
Matches 2059; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 55 CTGGGATCCCGAGCTCGGAGAGAGCA-CGGGCGGGCCCACTGCTGCTGCTGAGG 113
Db 1 CTGGGATCCCGAGCTCGGAGAGAGCA-CGGGCGGGCCCACTGCTGCTGCTGAGG 60
QY 114 CTCTGAGCCCGGCGGCGCCCGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAG 173
Db 61 CTCTGAGCCCGGCGGCGCCCGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 174 CTCTGAGCTGCTCTGAGGGGTTCCCTGTCAGAGAGAGAGAGAGAGAGAGAGAG 233
Db 121 CTCTGAGCTGCTCTGAGGGGTTCCCTGTCAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 234 TAGATACGAGAGTCCGTCAGAGAAAGAGAGTGAAGTGGGCGCCAGAGAGAGAG 293
Db 181 TAGATACGAGAGTCCGTCAGAGAAAGAGAGTGAAGTGGGCGCCAGAGAGAGAG 240
QY 294 CCGTCTGTTGCGCCCTGAGAGCCAGCTAAGTCTCAGATGTGCACTGCTGTGGCAG 353
Db 241 CCGTCTGTTGCGCCCTGAGAGCCAGCTAAGTCTCAGATGTGCACTGCTGTGGCAG 300
QY 354 CCGCGGCTGTGGGCGCTATGTCCTCTGAGAGCCAGAGAGAGAGAGAGAGAGAG 413

Db 301 CTTCCCGCTTGGGCGCTATGTCCTCTGAGAGCCGAGAGAGAGAGAGAGAGAGAG 360
QY 414 CCTGCACTAGCCCTTACAGAGAGAGTATACCTGAGAGAGTATACCCCTGCAAGAGCC 473
Db 361 CCTGCACTAGCCCTTACAGAGAGAGTATACCTGAGAGAGTATACCCCTGCAAGAGCC 420
QY 474 TGGCGGTGTGAGAGCCCTTACGCGGAGCTGCCCCCGCAAGAGATGTGCTGGGCCAG 533
Db 421 TGGCGGTGTGAGAGCCCTTATGCGCGAGCTGCCCCCGCAAGAGATGTGCTGGGCCAG 480
QY 534 AGTCTCTGCTGTATCCAGAGCTCTACGCTTTTTCATGAGAGAGAGAGAGAGAGAG 593
Db 481 AGTCTCTGCTGTATCCAGAGCTCTCTCTTACGCTTTTTCATGAGAGAGAGAGAG 540
QY 594 ACAGCTGTGTGAG 653
Db 541 ACAGCTGTGTGAG 600
QY 654 AGATGGCCACCGGCGCTGGGCGCATGTACACAGACAGCTGTGTCTGTGTATCTCAAG 713
Db 601 AGATGGCCACCGGCGCTGGGCGCATGTACACAGACAGCTGTGTCTGTGTATCTCAAG 660
QY 714 TGTGTGCTGTGTCTGT 773
Db 661 TGTGTGCTGTGTCTGT 720
QY 774 ACTCTGCTGT 833
Db 721 ACTCTGCTGT 780
QY 834 AGTGGAGCTGTAGATCTACGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 893
Db 781 AGTGGAGCTGTAGATCTACGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 894 GGAAGCTGTGGGCGT 953
Db 841 GGAAGCTGTGGGCGT 900
QY 954 AGCTGTCTGT 1013
Db 901 AGCTGTCTGT 960
QY 1014 CGGCGCTGTGCGGT 1073
Db 961 CGGCGCTGTGCGGT 1020
QY 1074 CAGCCAG 1133
Db 1021 CAGCCAG 1080
QY 1134 CCGGATCCCATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1193
Db 1081 CCGGATCCCATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
QY 1194 CCAAGGAG 1253
Db 1141 CCAAGGAG 1200
QY 1254 GGTCAAGCTGCAAG 1313
Db 1201 GGTCAAGCTGCAAG 1260
QY 1314 ACTGAGCCAAACCTTCAGTCTCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1373
Db 1261 ACTGAGCCAAACCTTCAGTCTCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1374 GGTGTGTATCATCTGCTTTTTCACATATGAGAGAGAGAGAGAGAGAGAGAGAG 1433
Db 1321 GGTGTGTATCATCTGCTTTTTCACATATGAGAGAGAGAGAGAGAGAGAGAGAG 1380
QY 1434 TGCAG 1493

```

Db      1381 TGCCAAAGCCCTGTTCTCGGTGCTGGAGATACAGACGAGCAAGCAAAAGACATATTCCT 1440
QY      1494 GCTCAAGAGATATCAAACTGGATCTTGAAGTGAACAACATTTTCCATGACCAATAGT 1553
Db      1441 GCTCAAGAGATATCAAACTGGATCTTGAAGTGAACAACATTTTCCATGACCAATAGT 1500
QY      1554 CACTGTCTACATCGGGTACACTTTGTACACAGTGTGGCTCCCACTAGTGTGGTCTAG 1613
Db      1501 CATGTCTACATCGGGTACACTTTGTACACAGTGTGGCTCCCACTAGTGTGGTCTAG 1560
QY      1614 GCACCTGTGTCCAGAGCAATCCCTTTACAAACAAACGAGTCCCTTTGTATCTGTAC 1673
Db      1561 GCACCTGTGTCCAGAGCAATCCCTTTACAAACAAACGAGTCCCTTTGTATCTGTAC 1620
QY      1674 CTTTGAAGAAAGGAGGATCCCTGTGCAAGGCTCCAGGCTCCCTCCGCAATC 1733
Db      1621 CTTTGAAGAAAGGAGGATCCCTGTGCAAGGCTCCAGGCTCCCTCCGCAATC 1680
QY      1734 AGGACCCAGCCAGCTCACTGTGGAACTGTGTCCAGCATCTGTCTGTATTA 1793
Db      1681 AGGACCCAGCCAGCTCACTGTGGAACTGTGTCCAGCATCTGTCTGTATTA 1740
QY      1794 AGAGATTCTCTTCCAGGCTTAAGCTGGGATTGGGCCAGAGATTAAGAACTAT 1853
Db      1741 AGAGATTCTCTTCCAGGCTTAAGCTGGGATTGGGCCAGAGATTAAGAACTAT 1800
QY      1854 GAGGCTAGTCTGTCTTAATCAAGACTGTCTGGAATGAGGCTCCAGGCTTCAACA 1913
Db      1801 GAGGCTAGTCTGTCTTAATCAAGACTGTCTGGAATGAGGCTCCAGGCTTCAACA 1860
QY      1914 TGGGGCTTCTGACCTGACCAAGGTTGAGGACAGAGATTAGCAGGGTCTGCTGTG 1973
Db      1861 TGGGGCTTCTGACCTGACCAAGGTTGAGGACAGAGATTAGCAGGGTCTGCTGTG 1920
QY      1974 GCCACCTGGAAGTCCAGGTGGGACTCTTCTGGGGAACCTTGGGTTCCAAATCCAG 2033
Db      1921 GCCACCTGGAAGTCCAGGTGGGACTCTTCTGGGGAACCTTGGGTTCCAAATCCAG 1980
QY      2034 TCATATCTAGGTTTGGATACCATGAGATGATGTTTACCTGTGCCAATAAGGAG 2093
Db      1981 TCATATCTAGGTTTGGATACCATGAGATGATGTTTACCTGTGCCAATAAGGAG 2040
QY      2094 AATTATGAATAAAAAAAAAAAAAA 2116
Db      2041 AATTATGAATAAAAAAAAAAAAAA 2063

```

```

RESULT 7
LOCUS   HS1103G7      153170 bp      DNA      linear      PRI 08-FEB-2001
DEFINITION
Human DNA sequence from clone RPS-1103G7 on chromosome 20p12.2-13.
Contains up to three novel genes, the gene for a novel protein
similar to mouse VMP, the gene for a novel protein kinase domain
containing protein similar to phosphoprotein C8W and rat NRPK, and
the SOX2 gene for SRY (sex-determining region Y)-box 22. Contains
five CpG islands, ESTs, STSs and GSSs, complete sequence.
ACCESSION
AL034548      GI:7263904
VERSION
AL034548.25
KEYWORDS
HTG, CpG island, NRPK, protein kinase, SOX2, SRY, VMP.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 153170)
AUTHORS
Blakey,S.
TITLE
Direct Submission
JOURNAL
Submitted (05-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
COMMENT
On Mar 19, 2000 this sequence version replaced gi:5541861.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the

```

FEATURES

variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/projects/c.elegans/wormpep> This sequence is the entire, insert of clone RPS-1103G7 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone, and the assembly was confirmed by restriction digest. This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr20> RPS-1103G7 is from the library RPT-5 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pCYPAC2.
Location/Qualifiers

```

1..153170
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="p12.2-13"
/clone="RPS-1103G7"
/clone_1fb="RPT-5"
1..61
/feature="Alu repeat: matches 2..62 of consensus"
79..129
/feature="L1M3 repeat: matches 6221..6271 of consensus"
130..173
/feature="22 copies 2 mer ag 79% conserved"
175..873
/feature="L1M3 repeat: matches 5519..6217 of consensus"
902..1109
/feature="L1M5 repeat: matches 7715..7921 of consensus"
1110..1480
/feature="TH1C repeat: matches 1..371 of consensus"
1502..1733
/feature="match: STS: Em:Z66755"
1734..1888
/feature="MLT1H repeat: matches 387..547 of consensus"
1919..2101
/feature="MLT1H repeat: matches 115..298 of consensus"
complement(join(2194..2243,126792..126871))
/feature="match: GSS: Em:AQ377604"
2699..2917
/feature="L1M4 repeat: matches 2875..3081 of consensus"
2938..3062
/feature="L1M4 repeat: matches 3120..3249 of consensus"
3136..3430
/feature="AluSg repeat: matches 1..295 of consensus"
3482..3777
/feature="AluSg repeat: matches 1..296 of consensus"
3811..4126
/feature="AluSx repeat: matches 1..312 of consensus"
4310..4608
/feature="L1M2 repeat: matches 5861..6167 of consensus"
4609..5026
/feature="WSTA repeat: matches 1..424 of consensus"
5027..5200
/feature="L1M2 repeat: matches 5684..5861 of consensus"
5201..5501

```

```
/note="AluSx repeat: matches 1. .307 of consensus"  
5502. .5727  
/note="L1MD2 repeat: matches 5451. .5664 of consensus"  
5798. .5865  
/note="34 copies 2 mer aa 66% conserved"  
6056. .6643  
/note="MER67C repeat: matches 122. .710 of consensus"  
6648. .6707  
/note="30 copies 2 mer ca 88% conserved"  
6710. .6824  
/note="MER50 repeat: matches 10. .134 of consensus"  
6825. .7124  
/note="AluSx repeat: matches 7. .306 of consensus"  
7126. .7415  
/note="L1M4A repeat: matches 5619. .5919 of consensus"  
7419. .7569  
/note="AluSg/x repeat: matches 148. .298 of consensus"  
7576. .7889  
/note="match: GSS: Em:AQ180303"  
/note="match: GSS: Em:AQ701808"  
complement(7590. .7894)  
7862. .8140  
/note="match: GSS: Em:AQ123077"  
8019. .8317  
/note="match: GSS: Em:AQ746749"  
8019. .8151  
/note="match: GSS: Em:AQ617736"  
8152. .8316  
/note="MER4B repeat: matches 416. .574 of consensus"  
8349. .8643  
/note="MER4D repeat: matches 362. .689 of consensus"  
8644. .8942  
/note="AluY repeat: matches 1. .299 of consensus"  
8943. .9113  
/note="MER4D repeat: matches 194. .362 of consensus"  
9113. .9201  
/note="MLT2PA repeat: matches 377. .450 of consensus"  
9226. .9311  
/note="L1MB4 repeat: matches 5775. .5862 of consensus"  
9312. .9651  
/note="MER1B repeat: matches 1. .337 of consensus"  
9653. .9750  
/note="L1MA5 repeat: matches 5856. .5947 of consensus"  
9751. .10047  
/note="AluSx repeat: matches 1. .298 of consensus"  
10048. .10369  
/note="L1MA5 repeat: matches 5947. .6300 of consensus"  
10488. .10529  
/note="21 copies 2 mer at 88% conserved"  
10532. .10653  
/note="L1MC4 repeat: matches 7673. .7800 of consensus"  
10687. .10752  
/note="33 copies 2 mer at 68% conserved"  
11681. .11742  
/note="31 copies 2 mer ta 71% conserved"  
13519. .13972  
/note="match: GSS: Em:AQ224309"  
14208. .14554  
/note="LTR16B repeat: matches 110. .462 of consensus"  
15554. .15872  
/note="L1MB1 repeat: matches 5615. .5910 of consensus"  
15873. .16174  
/note="AluSp repeat: matches 16. .313 of consensus"  
16175. .16382  
/note="L1MB1 repeat: matches 5400. .5615 of consensus"  
16383. .16423  
/note="L1ME repeat: matches 5893. .5933 of consensus"  
16595. .16894  
/note="AluY repeat: matches 1. .302 of consensus"  
16895. .17021  
/note="L1MA10 repeat: matches 6196. .6318 of consensus"  
17155. .17587  
/note="L1ME repeat: matches 5289. .5733 of consensus"
```

```
repeat_region 17587. .18157  
/note="L1M1 repeat: matches -1389. -.790 of consensus"  
repeat_region 18158. .18195  
/note="19 copies 2 mer tg 97% conserved"  
repeat_region 18197. .18230  
/note="AluSg/x repeat: matches 165. .197 of consensus"  
repeat_region 18231. .18823  
/note="MER52A repeat: matches 1. .672 of consensus"  
complement(18827. .19180)  
misc_feature /note="match: GSS: Em:AQ59338"  
19689. .19998  
repeat_region /note="AluSg repeat: matches 1. .309 of consensus"  
20287. .20543  
repeat_region /note="AluSx repeat: matches 2. .307 of consensus"  
20552. .20609  
repeat_region /note="MLT1J repeat: matches 255. .318 of consensus"  
20592. .20756  
repeat_region /note="MLT1E repeat: matches 256. .421 of consensus"  
20895. .21012  
repeat_region /note="L1P8a repeat: matches -1128. -.1008 of consensus"  
21009. .21168  
repeat_region /note="L1P8a repeat: matches -1537. -.1377 of consensus"  
21192. .21277  
repeat_region /note="L1MA7 repeat: matches 5864. .5949 of consensus"  
21280. .21627  
repeat_region /note="L2 repeat: matches 2250. .2617 of consensus"  
21628. .21884  
misc_feature /note="match: GSS: Em:B41180"  
21628. .21744  
/note="match: GSS: Em:AQ058160"
```

Query Match 64.6%; Score 1367; DB 9; Length 153170;
Best Local Similarity 98.9%; Pred. No. 3.2e-275;
Matches 1376; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```
QY 716 TGTGCTTGTCTTCTGCTGACCGCTGAGAGAGAAAGCTGTGTGAGAACTGTGAGAGC 775  
DB 150847 TGACCCCTTCTGTTTCTCCCAATGTCAGAGAAAGAGCTGTGTGAGAACTGTGAGAGC 150906  
QY 776 TCCTGCGTGCAGCTGGGCGGAGATATCCCTGTGGGACAGAGAGCTGTGACGCTTAC 835  
DB 150907 TCTGCGGTGCTGACCTGGGCGGAGATATTCCTGTGGGACAGAGAGCTGTGACGCTTAC 150966  
QY 836 GTGGAGCTGAGAACTAGCTCAAGGAGCTCATATCTGGGAGAGGCGGATGTCTGG 895  
DB 150967 GTGGGAGCTGAGAACTAGCTCAAGGAGCTCATATCTGGGAGAGGCGGATGTCTGG 151026  
QY 896 AGCTGGGCTGGGCTTCTTACCAATGTGGCGGCACTAACCCCTTCCAGAGACTCGAG 955  
DB 151027 AGCTGGGCTGGGCTTCTTACCAATGTGGCGGCACTAACCCCTTCCAGAGACTCGAG 151086  
QY 956 CCTGTCCGCTCTTGGGAGATCGCGGCGGAGCTTACGCTGTGAGAGCTTCTTCG 1015  
DB 151087 CTTGTCTGCTTCTTGGGAGATCGCGGCGGAGCTTACGCTGTGAGAGCTTCTTCG 151146  
QY 1016 GCCCTGCGCGCTGTCTGTGCTGTGCTCTTGTGTGGGAGCGAGCTGAAGGCTCA 1075  
DB 151147 GCCCTGCGCGCTGTCTGTGCTGTGCTCTTGTGTGGGAGCGAGCTGAAGGCTCA 151206  
QY 1076 GCCACAGGATCTCTTGCACCCCTGCTGAGCAGAGCCGATGCCCTTAAGCCCAAC 1135  
DB 151207 GCCACAGGATCTCTTGCACCCCTGCTGAGCAGAGCCGATGCCCTTAAGCCCAAC 151266  
QY 1136 CGATCCCATCTCTGGGAGGCTGCCAGAGTGTCTCCATATGATCGGGCTGTGAGAGCC 1195  
DB 151267 CGATCCCATCTCTGGGAGGCTGCCAGAGTGTCTCCATATGATCGGGCTGTGAGAGCC 151326  
QY 1196 AGGGAAGAGAGGAGAGCAGAGAGTGTCTGTATGCTAGACCAACCTACTACAGCC 1255  
DB 151327 AGGGAAGAGAGGAGAGCAGAGAGTGTCTGTATGCTAGACCAACCTACTACAGCC 151386  
QY 1256 TCAGCTGCCAAGTGAATTGAGTTGGGGGTAGCTCAAGCCTTCTTCGCTTGAAC 1315  
DB 1315 TCAGCTGCCAAGTGAATTGAGTTGGGGGTAGCTCAAGCCTTCTTCGCTTGAAC 1315
```

Db 151387 TCAGTCCCAACAGTGAATTGAGTTGGGGGAGTCCAGCTTCTCTGCTCTGAC 151446
 QY 1316 TGAGCCAAACCTTCAGTGCCTTCCAGAAAGGAGAAAGCCAGAGCTGTGTGAGTGTGC 1375
 Db 151447 TGAGCCAAACCTTCAGTGCCTTCCAGAAAGGAGAAAGCCAGAGCTGTGTGAGTGTGC 151506
 QY 1376 TGTGTACACATCTGCTTTGTTTCCACACATGCACTTCTGCTGGGTGCTTATCAGGTG 1435
 Db 151507 TGTGTACACATCTGCTTTGTTTCCACACATGCACTTCTGCTGGGTGCTTATCAGGTG 151566
 QY 1436 CCAAGCCCTGTTCTGCTGCTGGAGTACAGCAGTGAAGAAAGAGACATATTCCTGTC 1495
 Db 151567 CCAAGCCCTGTTCTGCTGCTGGAGTACAGCAGTGAAGAAAGAGACATATTCCTGTC 151626
 QY 1496 TCACAGAGATGACAACTGCGCATCTTGTAGCTGACAACTTTTCCATGACATATGCTCA 1555
 Db 151627 TCACAGAGATGACAACTGCGCATCTTGTAGCTGACAACTTTTCCATGACATATGCTCA 151686
 QY 1556 CTGTCTACCTGGGTACATTTGTACAGAGTCCGCTCCATGATGCTGCTCAGGC 1615
 Db 151687 CTGTCTACCTGGGTACATTTGTACAGAGTCCGCTCCATGATGCTGCTCAGGC 151746
 QY 1616 ACCTCTGCTCAAGACATCTCTTTCACAAACAAACCAAGCTGCTTGTATCTGTACT 1675
 Db 151747 ACCTCTGCTCAAGACATCTCTTTCACAAACAAACCAAGCTGCTTGTATCTGTACT 151806
 QY 1676 TTTTCAGAGAAAGGAGATATCCCTGTGCAAAAGCTCCAGGCTCTCCCTTGCATCTAG 1735
 Db 151807 TTTTCAGAGAAAGGAGATATCCCTGTGCAAAAGCTCCAGGCTCTCCCTTGCATCTAG 151866
 QY 1736 GACCCAGGCCAGCTCCTGCTGGGAATGTTGCCAGATCTCTGCTCTTGTATTAAG 1795
 Db 151867 GACCCAGGCCAGCTCCTGCTGGGAATGTTGCCAGATCTCTGCTCTTGTATTAAG 151926
 QY 1796 AGATTCTCTCTCCAGGCTTAAAGCTTGGGCAAGATTAAGATCAATCAACTATGA 1855
 Db 151927 AGATTCTCTCTCCAGGCTTAAAGCTTGGGCAAGATTAAGATCAATCAACTATGA 151986
 QY 1856 GGGTATGTTCTTGTCTTAACTGACAGCTGTTGGAATGAGGGTCCAGGCTGTCTCAACATG 1915
 Db 151987 GGGTATGTTCTTGTCTTAACTGACAGCTGTTGGAATGAGGGTCCAGGCTGTCTCAACATG 152046
 QY 1916 GGGCTTCTACCTGACACCAACCAAGTTGAGGGAAGATTAAGAGGGGTCTGTCTGTGAC 1975
 Db 152047 GGGCTTCTACCTGACACCAACCAAGTTGAGGGAAGATTAAGAGGGGTCTGTCTGTGAC 152106
 QY 1976 CACCTGGAAGTCCAGGTGGGACTCTTCTGGGGGACATTTGGGGTCCAAATCCAGGTC 2035
 Db 152107 CACCTGGAAGTCCAGGTGGGACTCTTCTGGGGGACATTTGGGGTCCAAATCCAGGTC 152166
 QY 2036 CATACTCTAGGTTTGGATTCATGAGATATGTTTACCTGTGCTTAATAAGAGAA 2095
 Db 152167 CATACTCTAGGTTTGGATTCATGAGATATGTTTACCTGTGCTTAATAAGAGAA 152226
 QY 2096 TTAATGAATTA 2106
 Db 152227 TTAATGAATTA 152237

RESULT 8
 AX166518 1077 bp DNA linear PAT 22-JUN-2001
 LOCUS AX166518
 DEFINITION Sequence 9 from Patent W00138503.
 ACCESSION AX166518
 VERSION AX166518.1 GI:14546863
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens (human)
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Piomman, G.D., Whyte, D., Manning, G.S., Sudarsanam, S.S., Martinez, R.,
 Planagan, P. and Clary, D.S.

TITLE Novel human protein kinases and protein kinase-like enzymes
 JOURNAL Patent: WO 0138503-A 9 31-MAY-2001;
 Sugen, Inc. (US)
 FEATURES Location/Qualifiers
 source 1..1077
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Query Match 50.7%; Score 1072.2; DB 6; Length 1077;
 Best Local Similarity 99.7%; Pred. No. 1.5e-213;
 Matches 1074; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 161 ATGGAGGCAACCCCTGCTGCTCTCTGAGGGTTCCTGTCAGAGAAAGCGTTGAG 220
 1 ATGGAGGCAACCCCTGCTGCTCTCTGAGGGTTCCTGTCAGAGAAAGCGTTGAG 60
 221 TTGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAAGAGCTCGAAATGCGCCAG 280
 61 TTGATGACAACTTAGATACCGAGCGTCCCGTCCAGAAAGAGCTCGAAATGCGCCAG 120
 281 CCCAGACTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 340
 121 CCCAGACTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 341 GCTGTGGCACTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 400
 181 GCTGTGGCACTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
 401 CGGGCTTACAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 460
 241 CGGGCTTACAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 461 GTTCAGAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 520
 301 GTTCAGAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 521 GCTGGGCTTACAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 580
 361 GCTGGGCTTACAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 581 CATGGGACATGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 640
 421 CATGGGACATGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 641 GTGCTTCTTCCGCAATGAGCCACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 700
 481 GTGCTTCTTCCGCAATGAGCCACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 701 CGTATCTCAAGCTGTGTGCTTGTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 760
 541 CGTATCTCAAGCTGTGTGCTTGTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 761 GAGAACCTGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 820
 601 GAGAACCTGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 821 GCTGTCCAGCCCTTACGTTGGGAGCTGAGATTAAGCTCAAGGCTCAATCTGCTGCTGCTGCTGCT 880
 661 GCTGTCCAGCCCTTACGTTGGGAGCTGAGATTAAGCTCAAGGCTCAATCTGCTGCTGCTGCTGCT 720
 721 GAGCCGATCTTGGAGCTGTGGGCTGTGGGCTCTTACCATGCTGGGCTGCTGCTGCTGCTGCTGCT 780
 941 TTCAGAGCTGAGAGCTGTCTGCTGCTTGTGGAAGATCCGCGCGGGGCTTACGCTTGG 1000
 781 TTCAGAGCTGAGAGCTGTCTGCTGCTTGTGGAAGATCCGCGCGGGGCTTACGCTTGG 840
 1001 CTGTCAGAGCTTCTGAGGCTGTGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1060
 841 CTGTCAGAGCTTCTGAGGCTGTGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900

QY 1061 GCTGAACGGCTCAGAGCCAGGCGATCCTCTGCGACCCCTGGCTGGGACAGACCCGATG 1120
DB 901 GCTGAACGGCTCAGAGCCAGGCGATCCTCTGCGACCCCTGGCTGGGACAGACCCGATG 960
QY 1121 CCCTTAGCCCAACCCGATCCATCTCTGGAGGCTGCCAGGCTGCTCCCTGATGACTG 1180
DB 961 CCCTTAGCCCAACCCGATCCATCTCTGGAGGCTGCCAGGCTGCTCCCTGATGACTG 1020
QY 1181 GGGCTGAGAGAGCCGAGGAG 1237
DB 1021 GGGCTGAGAGAGCCGAGGAG 1077

RESULT 9
AY335754 1077 bp mRNA linear SYN 15-OCT-2003
LOCUS
DEFINITION Synthetic construct Homo sapiens chromosome 20 open reading frame
ACCESSION AY335754
VERSION AY335754.1 GI:33304146
KEYWORDS FLI, CDNA.
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 1077)
AUTHORS Park,J., Rolfs,A., Hu,Y., Shen,B., Vannberg,F., Moreira,D., Kelley,T., Zuo,D., Raphael,J., Baqui,M., Jepson,D., Harlow,E., Labber,J., and Brizuela,L.
TITLE Cloning of human full-length CDS FLEXGene kinases in recombinational vector system
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1077)
AUTHORS Park,J., Rolfs,A., Hu,Y., Shen,B., Vannberg,F., Moreira,D., Kelley,T., Zuo,D., Raphael,J., Baqui,M., Jepson,D., Harlow,E., Labber,J., and Brizuela,L.
TITLE Direct Submission
JOURNAL Submitted (02-JUL-2003) Biological Chemistry and Molecular Pharmacology, Harvard Institute of Proteomics, 320 Charles St., Cambridge, MA 02141-2023, USA
COMMENT This CDS clone is a part of a collection of human full-length expression clones generated by Harvard Institute of Proteomics. Each CDS has been cloned without stop-codon (to allow fusion with C-terminal tag). The CDS has been directionally cloned using BD In-Fusion(TM) cloning system between the SalI and HindIII sites of the pDNR-Dual vector. Additional sequences in the clone: 'ACC' after SalI site and before 'ATG' to provide Kozak consensus sequence; 'GG' after last codon and before HindIII site to maintain reading frame.

FEATURES
source location/Qualifiers
1..1077
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/clone="FLH000134.01L"
/clone_id="MGC clone templates"
/lab_host="DH5alpha TI resistant"
/note="Vector: pDNR-Dual"
1..>1077
/gene="C20orf97"
1..>1077
/gene="C20orf97"
/note="Mutations: Stop->Leu"
/transl_table=1
/product="chromosome 20 open reading frame 97"
/protein_id="AA02581.1"
/db_xref="GI:33304147"
/translation="MKNPPLAAGSLSRKRLDLDNLTERPVQKRAAGSGPQRLPFCULPLSPDAVATASRLGPLYLLEPEGGRAYQALHCPTEGYCKYPVQBALVLEIYARLPKHVARLPTEVLGATQLEAFRTGHGMSLIVSRRIPEEAAVLFROMATLALHCHQHGVLRLDKLCRFVFAERKRKLVLENLDSCLTGPDSIMDKACPAVVGPEILSSRASYSKADVMSLGVALTFLMLAGHYPPDSEPVLLFGKIRRG

ORIGIN
Query Match 50.6%; Score 1070.6; DB 12; Length 1077;
Best Local Similarity 99.6%; Pred. No. 3,2e-213;
Matches 1073; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
AYALPAGSAPARCLVRCILREPAERLTATGILHPLRLQDPMPLAPTRSHLWAAQ
VVPDGLGIDEAREEBEDREVLYGL"
QY 161 ATGCGAGCCACCCCTCTGCGTCTCTGCGGAGTTCCTCTCCAGAAAGAGCGTTGAG 220
DB 1 ATGCGAGCCACCCCTCTGCGTCTCTGCGGAGTTCCTCTCCAGAAAGAGCGTTGAG 60
QY 221 TTGATGACACTTGAATACGAGCGTCCGTTCCAGAAAAGAGCTGAAGTGGCCCAAG 280
DB 61 TTGATGACACTTGAATACGAGCGTCCGTTCCAGAAAAGAGCTGAAGTGGCCCAAG 120
QY 281 CCGAGACTGCCCCCTGCGCTGTGTCGCCCTGAGCCCACTTACATGATGTGCAACT 340
DB 121 CCGAGACTGCCCCCTGCGCTGTGTCGCCCTGAGCCCACTTACATGATGTGCAACT 180
QY 341 GCTGTGACACTGCTCCGCTCTTGTGGCCCTATATGTCCTGAGAGCCGAGAGGCGG 400
DB 181 GCTGTGACACTGCTCCGCTCTTGTGGCCCTATATGTCCTGAGAGCCGAGAGGCGG 240
QY 401 CGGCTTACAGAGGCTTGTGACTGACCTTACAGAGCACTGATATACCTGCAAGGTGACCC 460
DB 241 CGGCTTACAGAGGCTTGTGACTGACCTTACAGAGCACTGATATACCTGCAAGGTGACCC 300
QY 461 GTCCAGAGAGCGCTTGGCGGTGTGAGAGCCCTACAGCGCGCTGCCCCGCAAGCATGTG 520
DB 301 GTCCAGAGAGCGCTTGGCGGTGTGAGAGCCCTACAGCGCGCTGCCCCGCAAGCATGTG 360
QY 521 GCTCGGCCACTGAGGCTCTGGCTGTATACCAAGCTCTTACAGCGCTTTTCTACTCGGACC 580
DB 361 GCTCGGCCACTGAGGCTCTGGCTGTATACCAAGCTCTTACAGCGCTTTTCTACTCGGACC 420
QY 581 CATGGGACATGACACAGCTGTGTGGAAGCGGCCACCGATATCCCTGAGGCTTGGCTGCC 640
DB 421 CATGGGACATGACACAGCTGTGTGGAAGCGGCCACCGATATCCCTGAGGCTTGGCTGCC 480
QY 641 GTCCTTTCGCGCAGATGGCCACCGGCTTGGCGCACTGTACACAGACAGGTTGGTCTTG 700
DB 481 GTCCTTTCGCGCAGATGGCCACCGGCTTGGCGCACTGTACACAGACAGGTTGGTCTTG 540
QY 701 CGTATCTCAAGCTGTGTGCTTTGCTTCTGCTGACGCTGAGAGAAAGCTGTGCTG 760
DB 541 CGTATCTCAAGCTGTGTGCTTTGCTTCTGCTGACGCTGAGAGAAAGCTGTGCTG 600
QY 761 GAGAACTGAGAGACTCTGCTGTGCTGACTGAGGCGCATGATATTCCTGTGGACAAGAC 820
DB 601 GAGAACTGAGAGACTCTGCTGTGCTGACTGAGGCGCATGATATTCCTGTGGACAAGAC 660
QY 821 GCGTGGCCAGGCTACGTGGGACCTGAGATATGAGCTACAGGGCTCTATCTCGGGCAAG 880
DB 661 GCGTGGCCAGGCTACGTGGGACCTGAGATATGAGCTACAGGGCTCTATCTCGGGCAAG 720
QY 881 GCAAGCCGATGCTGAGAGCTGGGCGCTCTTCAACATGCTGAGCGGCACTAACCCC 940
DB 721 GCAAGCCGATGCTGAGAGCTGGGCGCTCTTCAACATGCTGAGCGGCACTAACCCC 780
QY 941 TTCCAGAGCTGAGAGCTGCTCTCTCTTTCGGAAGATCCGCGGAGGCTTACGCTTGTG 1000
DB 781 TTCCAGAGCTGAGAGCTGCTCTCTCTTTCGGAAGATCCGCGGAGGCTTACGCTTGTG 840
QY 1001 CCTGCAAGGCTCTTGGGCGCTGCGGCTGTCTGTGCTGCTGCTCTTCTGTCGGAGCA 1060
DB 841 CCTGCAAGGCTCTTGGGCGCTGCGGCTGTCTGTGCTGCTGCTCTTCTGTCGGAGCA 900
QY 1061 GCTGAAGGGCTCAGAGCAAGGAGATCTCTCTGAGCCCTGAGCGGCGACAGAGCCGATG 1120
DB 901 GCTGAAGGGCTCAGAGCAAGGAGATCTCTCTGAGCCCTGAGCGGCGACAGAGCCGATG 960
QY 1121 CCCTTAGCCCAACCCGATCCATCTCTGGAGGCTGCCAGAGTGTCTCTGATGAGACTG 1180

QY	755	GTGCTGGAGAACCTCGAGGAACTCCTGGGTGTGACTGGGCGAATGATTCCTGTGGAC	814
Db	601	GTCTGTGAGAACTCTGGAGAACTCTGTGTGTGACTGGGCGAATGATTCCTGTGGAC	660
QY	815	AAGCAGCGGTGTCCAGCCTTACGTGAGAACTGAGATTACTCAGCTCACGGGCTCATTA	874
Db	661	AAGCAGCGGTGTCCAGCCTTACGTGAGAACTGAGATTACTCAGCTCACGGGCTCATTA	720
QY	875	GGCAAGGAGCCGATGTCGTGAGGCTTGAGGCGCTTCAACATGCTGGCGGAC	934
Db	721	GGCAAGGAGCCGATGTCGTGAGGCTTGAGGCGCTTCAACATGCTGGCGGAC	780
QY	935	TACCCCTTCCAGAACTGTGAGGCTGTGTGCTCTTCGCAAGATCCGCGGGGCTTAC	994
Db	781	TACCCCTTCCAGAACTGTGAGGCTGTGTGCTCTTCGCAAGATCCGCGGGGCTTAC	840
QY	995	GCCTTGTGCTGAGGACCTCTGGGCCCCCGGCTGTGAGTGTGGCTGCTTGTGTCGG	1054
Db	841	GCCTTGTGCTGAGGACCTCTGGGCCCCCGGCTGTGAGTGTGGCTGCTTGTGTCGG	900
QY	1055	GAGCCAGCTGAAGGAGCTCAGAGCCACAGGCAATCTCTCTGCACCCCTGTGGTGGACAGGAC	1114
Db	901	GAGCCAGCTGAAGGAGCTCAGAGCCACAGGCAATCTCTCTGCACCCCTGTGGTGGACAGGAC	960
QY	1115	CCGATGCCCCCTTAAAGCCCCCAACCCGATCCGATCTCTGGGAGGGTCCAGAGTGA	1174
Db	961	CCGATGCCCCCTTAAAGCCCCCAACCCGATCCGATCTCTGGGAGGGTCCAGAGTGA	1020
QY	1175	GGACTGGGGCTGAGCGAAGCCAGGGAGAGGAGGAGACAGAGAAATGTTCCTGTATGGC	1234
Db	1021	GGACTGGGGCTGAGCGAAGCCAGGGAGAGGAGGAGACAGAGAAATGTTCCTGTATGGC	1080
QY	1235	TAG 1237	
Db	1081	TAG 1083	

RESULT 13			
AF250311			
LOCUS	AF250311	1083 bp	mRNA
DEFINITION	Homo sapiens SKIP3 mRNA, complete cds.		linear
ACCESSION	AF250311		
VERSION	AF250311.1	GI:14276268	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 1083)		
AUTHORS	Kiss-Toth,E., Wyllie,D.H., Qvarnstrom,E.E. and Dower,S.K.		
TITLE	Identification of pro-inflammatory cytokine signalling network components by transcription expression screening		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1083)		
AUTHORS	Kiss-Toth,E., Wyllie,D.H., Qvarnstrom,E.E. and Dower,S.K.		
TITLE	Direct Submission		
JOURNAL	Submitted (29-MAR-2000) Division of Molecular & Genetic Medicine, University of Sheffield, Royal Hallamshire Hospital, Floor M, Sheffield, S10-2TF, UK		
FEATURES	Location/Qualifiers		
source	1..1083		

CDS

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..1083
/codon_start=1
/product="SKIP3"
/protein_id="AAK58175.1"
/db_xref="GI:44276269"
/translation="MRRTPLAAGSLSRKKLELDNDIPERPYOKARSGDPPRLP
ECILPLSPFRPADATAVATASLSGYVLLPEEGRAVOALHCTGTEYCKIIPYQ
EALVAVLAAVPHKVNAPRILVLTGLTGLRFLRTHGDMALSVSRRAHLPPEPDA
LTFOMALAAHQHQLVTRDILKTRFPFAADIRKKLLVETNELSDSCVLTGPDSD

```

ORIGIN

Query Match	49.8%	Score 1053;	DB 9;	length 1083;
Best Local Similarity	99.0%;	Pred. No. 1.5e-209;		
Matches 1072;	Conservative	0;	Mismatches 5;	Indels 6;
				Gaps 1;

WKHACPAVYGPPELLSSPA5YSGKADWSLGVAF7MLAGHYPPQDSEVLLFGKLR
 RGVAYLPAGASAPARCLYRCLLRBEAEERLTAT5ILHPWLRQDPWPLAPTRSHLWEA
 AQVVPDGLGDEAREEEDREVLVYQ³

QY	161	ATGGAGCCACCCCTCGAGGTGCTCTCGAGGGTTCCTGTGTCAGAGAAAGCGATTGAG	220
Db	1	ATGGAGCCACCCCTCGAGGTGCTCTCGAGGGTTCCTGTGTCAGAGAAAGCGATTGAG	60
QY	221	TTGGATGACAACTTAAGATACCGAGCGTCCGCTCCAGAAAGAGGTCCGAGCGGCGCCAG	280
Db	61	TTGGATGACAACTTAAGATACCGAGCGTCCGCTCCAGAAAGAGGTCCGAGCGGCGCCAG	120
QY	281	CCCGAGTCGGCCCCCTGCTGTTCGCTCCCTGAGGCCACTTACTGTCTCCAGATCTGCAACT	340
Db	121	CCCGAGTCGGCCCCCTGCTGTTCGCTCCCTGAGGCCACTTACTGTCTCCAGATCTGCAACT	180
QY	341	GCTGTGGCCACTGTCCTCCCGTCTTGAGGCCCTATGTCTCTTGAGGCCGAGGAGGCGGG	400
Db	181	GCTGTGGCCACTGTCCTCCCGTCTTGAGGCCCTATGTCTCTTGAGGCCGAGGAGGCGGG	240
QY	401	CGGCGCTTACCAAGGCGCTTGCACTGCTCCCTTACAGGCACTAGATTTACTGCAAGGTATACCC	460
Db	241	CGGCGCTTACCAAGGCGCTTGCACTGCTCCCTTACAGGCACTAGATTTACTGCAAGGTATACCC	300
QY	461	GTCAGAGAAAGCCCTGGCGCGTGTGGAGCGCTTACGCGCGGCGTCCCGCCGACAAACATGTG	520
Db	301	GTCAGAGAAAGCCCTGGCGCGTGTGGAGCGCTTACGCGCGGCGTCCCGCCGACAAACATGTG	360
QY	521	GCTCGGCGCACTGAGGTCTGTGACTGTGATACCCAGCTCTCTTACGCTTTTTCACTCGAGCC	580
Db	361	GCTCGGCGCACTGAGGTCTGTGACTGTGATACCCAGCTCTCTTACGCTTTTTCACTCGAGCC	420
QY	581	CATGGGGACATGACACAGCTGTGTGTGCGAAACGCGCACCGATTCCTTGAGCTGAGAGTCC	640
Db	421	CATGGGGACATGACACAGCTGTGTGTGCGAAACGCGCACCGATTCCTTGAGCTGAGAGTCC	480
QY	641	GTGCTCTTCCGCGCAGATATGCGCACCGCCTGGCGGCACTGTACACAGCACCGATCTGTCTGTG	700
Db	481	GTGCTCTTCCGCGCAGATATGCGCACCGCCTGGCGGCACTGTACACAGCACCGATCTGTCTGTG	540
QY	701	CGTGAATCTCAAGCTGTGTGTGCTTTGTCTTGC-----TGAACGCTGAGAGGAGAAAGCTG	754
Db	541	CGTGAATCTCAAGCTGTGTGTGCTTTGTCTTGC-----TGAACGCTGAGAGGAGAAAGCTG	600
QY	755	GTGCTGGAACCTTGAGAGACTCTCTGCGTGTGACTGTGGGCGAAGATGATTCCTGTGTGGAC	814
Db	601	GTGCTGGAACCTTGAGAGACTCTCTGCGTGTGACTGTGGGCGAAGATGATTCCTGTGTGGAC	660
QY	815	AAGCACGCGTGTCCAGACTTACGTGGAGCTTAGATATCTACAGTCTACCGGAGCTCATATCTCG	874
Db	661	AAGCACGCGTGTCCAGACTTACGTGGAGCTTAGATATCTACAGTCTACCGGAGCTCATATCTCG	720
QY	875	GCGAAGCGAGCGATGTCTGGAGCTGGGCGGTGGGACTTTACCCATGCTGTGACCGGCGCAC	934
Db	721	GCGAAGCGAGCGATGTCTGGAGCTGGGCGGTGGGACTTTACCCATGCTGTGACCGGCGCGCAC	780
QY	935	TACCCCTTTCAGAGACTCGGAGCTGTCTGTCTTTCGGGCAAGATTCGCGCGGGGGGCTTAC	994
Db	781	TACCCCTTTCAGAGACTCGGAGCTGTCTGTCTTTCGGGCAAGATTCGCGCGGGGGGCTTAC	840
QY	995	GCGTTGCTGAGAGGCTCTGGGCGCTTGGCGCGGTGTGATTTGGTGGCTCTTGTGTGG	1054
Db	841	GCGTTGCTGAGAGGCTCTGGGCGCTTGGCGCGGTGTGATTTGGTGGCTCTTGTGTGG	900
QY	1055	GAGCCAGCTGAACGGCTCTACAGCCACAGGACATCTCTTGCAACCCCTGTGCTGACAGGAC	1114
Db	901	GAGCCAGCTGAACGGCTCTACAGCCACAGGACATCTCTTGCAACCCCTGTGCTGACAGGAC	960

QY 1115 CCGATGCCCTTAAAGCCCAACCCGATCCATCTCTGGAGGCTGCCAGGAGTCCCTGAT 1174
DB 961 CGATATCCCTTACGCCCCCAACCCGATCCATCTCTGGAGGCTGCCAGGAGTCCCTGAT 1020
QY 1175 GGAATGGGCTGTGACGAGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1234
DB 1021 GGAATGGGCTGTGACGAGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1235 TAG 1237
DB 1081 TAG 1083

RESULT 14
AX224734 972 bp DNA linear PAT 10-SBP-2001
LOCUS AX224734
DEFINITION Sequence 21 from Patent WO0160991.
ACCESSION AX224734
VERSION AX224734.1 GI:15554836
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1
Tang, Y. T., Buford, N., Gandhi, A. R., Patterson, C., Khan, F. A., Yue, H.,
Hafila, A., Shih, L. L., Trilouley, C. M., Yao, M. G., Burxilli, J. D.,
Marus, G. A., Zingler, K. A., Lu, D. A., Bandman, O., Policky, D. L.,
Griffin, J. A., Thornton, M., Nguyen, D. B. and Walsh, R. T.,
Human Kinases
TITLE Patient: WO 0160991-A 21.23-AUG-2001;
JOURNAL Incyte Genomics, Inc. (US)
FEATURES
source location/Qualifiers
1..972
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7472328CB1"

ORIGIN
Query Match 38.0%; Score 803.8; DB 6; Length 972;
Best Local Similarity 87.7%; Pred. No. 1.8e-157;
Matches 945; Conservative 0; Mismatches 27; Indels 105; Gaps 2;
QY 161 ATGCGAGCCACCCCTTGGCTGCTCTCTGGGGTTCCCTGTCAGAGAGAGCGATTGAG 220
DB 1 ATGCGAGCCACCCCTTGGCTGCTCTCTGGGGTTCCCTGTCAGAGAGAGCGATTGAG 60
QY 221 TTGATGACAACTTATGATACGAGAGCGTCCCGTCAGAGAGAGAGAGAGAGAGAGAGAG 280
DB 61 TTGATGACAACTTATGATACGAGAGCGTCCCGTCAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 281 CCGAGACTGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 340
DB 121 CCGAGACTGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 341 GCTGTGGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 400
DB 181 GCTGTGGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 401 CGGGGCTTACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 460
DB 241 CGGGGCTTACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 461 GTTCCAGAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 520
DB 301 GTTCCAGAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 521 GCTGTGGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 580
DB 361 GCTGTGGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 581 CATGGAGCATGCAAGGCTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 640

DB 421 CATGGAGCATGCAAGGCTGTAGG----- 446
QY 641 GTGCTTCCGCGAATGGCCACCCGCCCTGGGCACTGTCAACAGACAGGTCTGATCTG 700
DB 447 -----GCAACGCCCTGTGTGCACTGTGACAG----- 474
QY 701 CGTATCTCAAGCTGTGTGTGCTTTGTCTTGTGACCCGTGAGAGAGAGAGAGAGAGAGAG 760
DB 475 -----ACACGAG 495
QY 761 GAGAACCTTGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 820
DB 496 GAGAACCTTGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 555
QY 821 GGCTGCCAGGCTTACGTTGGGAGCCTGAGATACCTAGCTACGAGGCTTCAATCTGGGCAAG 880
DB 556 GGCTGCCAGGCTTACGTTGGGAGCCTGAGATACCTAGCTACGAGGCTTCAATCTGGGCAAG 615
QY 881 GCAGCCGATGTCTGAGAGCTTGGGCGTGGGCTCTTCAACCATGCTGACCGGCACTACCC 940
DB 616 GCAGCCGATGTCTGAGAGCTTGGGCGTGGGCTCTTCAACCATGCTGACCGGCACTACCC 675
QY 941 TTCCAGAGACTGAGAGCTGTCTGCTGCTTTCGCAAGATCCGCGCGGAGCTTACGCTTG 1000
DB 676 TTCCAGAGACTGAGAGCTGTCTGCTGCTTTCGCAAGATCCGCGCGGAGCTTACGCTTG 735
QY 1001 CCTGAGAGGCTCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1060
DB 736 CCTGAGAGGCTCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 795
QY 1061 GCTGAACGGCTTACAG 1120
DB 796 GCTGAACGGCTTACAG 855
QY 1121 CCTTATGCCCCCAACCCGATCCCATCTTGTGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1180
DB 856 CCTTATGCCCCCAACCCGATCCCATCTTGTGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 915
QY 1181 GGGCTGAG 1237
DB 916 GGGCTGAG 972

RESULT 15
AX364906 1076 bp DNA linear PAT 15-FEB-2002
LOCUS AX364906
DEFINITION Sequence 57 from Patent WO0206315.
ACCESSION AX364906
VERSION AX364906.1 GI:18696795
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1
Mintz, L., Freilich, S. and Bernstein, J.
Novel nucleic acid and amino acid sequences
TITLE Patent: WO 0206315-A 57 24-JAN-2002;
JOURNAL Compugen Ltd. (IL)
FEATURES
source location/Qualifiers
1..1076
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 33.3%; Score 716.6; DB 6; Length 1076;
Best Local Similarity 98.0%; Pred. No. 3e-139;
Matches 753; Conservative 3; Mismatches 8; Indels 4; Gaps 3;
QY 27 CTAGAGACCCGGGAGGCTGAGGCTGGAGTCCGAGCTCGGAGAGAGAGAGAGAGAGAGAG 86

```
Db      196 CTAGAGCCCGGCGAGGCGTGGAGCTGGGCTGGGATCCCGAGCTCGGCAGAGCGCA-GNG 254
QY      87  CCGGCCCACTTGTGTGTGCTGAGGCTCTGAGCCCGGCGGGCGGCCCACTCG 146
Db      255 CCGGCCCACTGTGTGTGCTGAGGCTCTGAGCCCGGCGGGCGGCCCACTCG 314
QY      147 AACGACGGGGCGAGATGGAGCAACCCCTGCTGCTCTGCGGGTTCCCTGTCAAGGA 206
Db      315 AACGACGGGGCGAGATGGAGCAACCCCTGCTGCTCTGCGGGTTCCCTGTCAAGGA 374
QY      207 AGAAGCGGTTGGAGTTGAGTGAACAATTAGATACGAGCGTCCCGTCCAGAAACGAGCTC 266
Db      375 AGAAGCGGTTGGAGTTGAGTGAACAATTAGATACGAGCGTCCCGTCCAGAAACGAGCTC 434
QY      267 GAAGTGGGCCCAAGCCAGACTGCCCCCTGCTGTGTCCTTGAAGCCCACTACTGCTC 326
Db      435 GAAGTGGGCCCAAGCCAGACTGCCCCCTGCTGTGTCCTTGAAGCCCACTACTGCTC 494
QY      327 CAGATCTGTGCAACTGTGTGGGCACTGCGCTCCCGTCTTGGGCGCTATGTCTCTGAGGC 386
Db      495 CAGATCTGTGCAACTGTGTGGGCACTGCGCTCCCGTCTTGGGCGCTATGTCTCTGAGGC 554
QY      387 CCGAGAGGGCGGGCGGGCCTTACAGAGCCCTGCACTGCCCCCTACAGGCACTGATATCCT 446
Db      555 CCGAGAGGGCGGGCGGGCCTTACAGAGCCCTGCACTGCCCCCTACAGGCACTGATATCCT 614
QY      447 GCAAGGTGTACCCCGTCCAGGAAGCCCTGCGGTGCTGAGGCCCTACGGCGGGCTGCC 506
Db      615 GCAAGGTGTACCCCGTCCAGGAAGCCCTGCGGTGCTGAGGCCCTATGCGGGCTGCC 674
QY      507 CGCACAAGCATGTGTGGCTCGGGCCCACTGAGGTCTGCTGTACCCAGCTCTCTACGCT 566
Db      675 CGCACAAGCATGTGTGGCTCGGGCCCACTGAGGTCTGCTGTACCCAGCTCTCTACGCT 734
QY      567 TTTTCACTCGGACCCATGGGGAATGACAGAGCTGTGTGGAAAGCGCAACCGTATCCCTG 626
Db      735 TTTTCACTCGGACCCATGGGGAATGACAGAGCTGTGTGGAAAGCGCAACCGTATCCCTG 794
QY      627 AGCTGAGGCTGCGGTCTTTCGCGAATGGGCCACCGCCCTGCGCACTGTCAACAGC 686
Db      795 AGCTGAGGCTGCGGTCTTTCGCGAATGGGCCACCGCCCTGCGCACTGTCAACAGC 854
QY      687 ACGGTCTGTGCTGCGGATCTCAAGCTGTGTGCTTTGTCTTGCTGAACCGTGAAGGA 746
Db      855 ACGGTCTGTGCTGCGGATCTCAAGCTGTGTGCTTTGTCTTGCTGAACCGTGAAGG- 913
QY      747 AGAAGCTGTGTGAGAACTGAGAGACTCTGCGTGTGACTGAGC 794
Db      914 AGAAGCTGTGTGAGAA--CTGAGAGACTCTGCGTGTGACTGAGC 959
```

Search completed: August 29, 2004, 10:31:30
Job time : 8351 secs